

# **aer**

**THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH**

**VOLUME XXV    NUMBER 1    MARCH 1979**

**PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    ●    EDMONTON**



# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

Published quarterly in March, June, September, December by the  
Faculty of Education, The University of Alberta

## CONSULTING EDITORS

J. Britton  
*University of London*

J. Calam  
*The University of British Columbia*

M. Connelly  
*The Ontario Institute for  
Studies in Education*

K. De Clerck  
*State University of Ghent*

R. N. Evans  
*University of Illinois at  
Urbana-Champaign*

R. H. Farquhar  
*University of Saskatchewan*

E. Gagné  
*University of Ottawa*

G. Harman  
*The Australian National University*

S. Hunka  
*The University of Alberta*

J. W. G. Ivany  
*Simon Fraser University*

D. A. MacIver  
*University of New Brunswick*

L. D. Nelson  
*The University of Alberta*

W. C. Nesbit  
*Memorial University of Newfoundland*

E. Pedersen  
*McGill University*

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

## FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith  
T. P. Atkinson  
N. C. Bhattacharya  
W. T. Fagan

H. W. Hodysh  
W. W. Laing  
R. G. Martin

E. Miklos (Chairman)  
J. W. Osborne  
C. H. Preitz

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

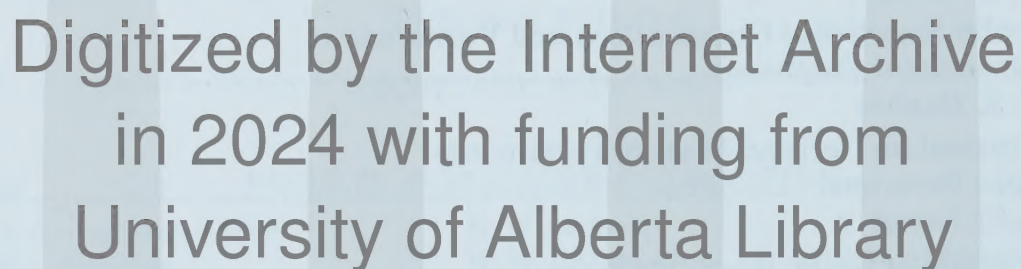
Volume XXV, Number 1

March, 1979

## CONTENTS

Guest Editorial: The Child and Educational Research For Whose Benefit? .....	1
<i>D. K. Kieren</i>	
Curricular Knowledge Organization and Variations in Instructional Emphases .....	4
<i>A. S. Hughes</i>	
The Curriculum Decision-Making Preferences of School Personnel .....	20
<i>J. H. Young</i>	
The Disturbing Child: He or She? .....	30
<i>L. Schlosser and B. Algozzine</i>	
Pupil Control Ideology: Comparative Perspectives— United States and Canada .....	37
<i>A. Hamalian</i>	
The Effects of Role Exchange Questioning on Empathetic Perceptiveness .....	48
<i>J. W. Kehoe and C. Ungerleider</i>	
BOOK REVIEWS	
<i>Register of Research into Higher Education in Western Europe 1974-7</i> .....	53
Reviewed by D. R. Pugh	
<i>Teaching Concepts: An Instructional Design Guide,</i> J. David Merrill and Robert D. Tennyson .....	54
Reviewed by David Baine	
<i>Assessing Students: How Shall We Know Them?</i> , Derek Rowntree .....	56
Reviewed by W. W. Laing	
<i>Children in English Canadian Society: Framing the Twentieth Century Consensus</i> , Neil Sutherland .....	57
Reviewed by J. S. Hardy	





Digitized by the Internet Archive  
in 2024 with funding from  
University of Alberta Library

[https://archive.org/details/ajer\\_1979](https://archive.org/details/ajer_1979)



## *Guest Editorial*

# The Child and Educational Research — For Whose Benefit?

Dianne K. Kieren

*Associate Professor and Division Chairperson  
Division of Family Studies, Faculty of Home Economics  
The University of Alberta*

1979 marks the occasion for nations to renew and reaffirm concern for the present and future development of their children. When the General Assembly of the United Nations passed a resolution declaring 1979 the International Year of the Child, they outlined the following major aims:

To encourage all countries, rich and poor, to review their programmes for the promotion of the well-being of children, and to mobilize support for national and local action programmes according to each country's conditions, needs and priorities;

To heighten awareness of children's special needs among decision-makers and the public;

To promote recognition of the vital link between programmes for children on the one hand, and economic and social progress on the other;

To spur specific, practical measures — with achievable goals — to benefit children, in both the short and long term on the national level. (United Nations, no date specified.)

As these aims indicate, the year is intended to stimulate thought and action toward developing each nation's most valuable resource — its children. Special attention is encouraged for the many children with special needs: handicapped children, slum children, abused children, orphans, refugee children, children exposed to life in families with limited personal and physical resources, and the many children suffering from malnutrition. The brochure circulated by the United Nations further states:

While programmes of national and local action remain the focal point of the year, International Year of the Child should also stimulate new research on the needs of children and the widespread gathering and dissemination of research on child related issues.

This call for research on children and their needs leads to the central



query of this editorial — research for whose benefit? Will International Year of the Child yield increased funding for child-related researchers and programmers or will it yield direct benefits to children? While the answer to this question may seem obvious to some, it is apparent that the significance of research activities generated during International Year of the Child, or any year for that matter, must be assessed at several levels: significance for researchers, for practitioners dealing with children, for society as a whole, and for children and their unique needs. One cannot assume that, because research is conducted about children, the data so gathered will immediately benefit them. In the strictest sense, the researcher is removed from the arena where his or her findings might have greatest impact — the home, the school, and the social service agency. In addition, a single study may provide only one piece of the large puzzle of human growth and interaction. Research is a time-consuming effort which demands special skills; thus the application of findings often are not immediately evident, even to those who understand the research. This being the case, why conduct research at all? Does research really make a difference in the day-to-day life and activities of children?

In order to answer this question, readers and contributors to this journal should give thought to how the relationship between researchers and practitioners in education might be strengthened. The gulf between the researcher who collects the data and the practitioner or consumer who ultimately applies the new information has continued to widen. Researchers are viewed as operating in sterile ivory towers, far from the mainstream of the problems which form the basis of their research. Practitioners looking for new insights in dealing with children turn from the research journals pleading that the studies therein are frequently irrelevant, that they focus on part of rather than on the whole child, and that they introduce unnecessary complexities in data collection and reporting. Researchers retort by claiming that practitioners lack motivation to maintain levels of scientific rigour and fail to apply the information with sufficient attention to the conditions of the research. The arguments could go on and on, but the issue remains that the effect of research on the daily lives of children depends upon a shared responsibility between researcher and practitioner. For the child living in a family in which he is constantly abused, it is of little importance to know that there are hundreds of studies going on at the present time attempting to determine the interactional and personality correlates of such behaviour. To the child experiencing a learning disability, it is irrelevant that numerous studies are being conducted at this very moment which might assist his teacher in developing techniques to meet his educational needs. The time lag between the collection of data by the researcher and the ultimate translation of the data to effective practice is far too lengthy. It is an increasing concern of researchers and practitioners alike that this time lag be shortened and that the applications of research data be ones that somehow enrich the human potential of our children. An increase in the benefits of research to children appears to depend on changed behaviour of researchers and practitioners alike.

Three areas are of particular concern to researchers and educational practitioners. These centre around the selection of research areas, the interpretation of research findings, and the application of these findings.



Selecting research topics can be the sole responsibility of the researcher or it can be a joint effort of researchers and practitioners. The potential advantages of cooperative work are many. Increased identification of problematic situations worthy of study, establishment of rapport between the two groups, and clearer conceptualization of the process of formulating research questions from practical concerns are illustrative of the benefits of closer work between researchers and practitioners. Joint involvement often leads to increased commitment to the goals of both research and practice.

With regard to the interpretation and application of research findings, researchers often worry that practitioners do not understand the research process well enough to appreciate the limits of generalizability, and thus become frustrated in attempting to apply the results of a single piece of research rather than fitting this work into a composite picture. Practitioners must come to realize the value of deriving generalizations or propositions from several studies of the same phenomenon. Researchers, however, have a responsibility to encourage this behaviour by replicating more studies and increasing the number of review articles which have as their main objective the derivation of such propositions. Just as a researcher might be challenged to test the propositions posed in such a review, so might the practitioner be led to refine them into useable teaching strategies.

International Year of the Child provides the appropriate stimulus to enlist the joint involvement of researchers, teachers, and parents to create environments which truly benefit children. A healthy dialogue between researchers and practitioners will contribute much to the achievement of this goal.

#### *Reference*

United Nations. *International Year of the Child*. Mimeographed brochure. New York: International Year of the Child Secretariat, no date specified.



ANDREW S. HUGHES

*Atlantic Institute of Education*

## Curricular Knowledge Organization and Variations in Instructional Emphases

*This study was designed to examine the relationship between the use of collection and integrated curricular knowledge organizations and the relative emphasis given selected aspects of instruction. Data from five independently conducted case studies were collected by trained observers employing a classroom observation category system designed for the purposes of the study. The findings suggest that "instructional intent" and the conditions of "knowing that" are independent of the curricular knowledge organization in use, while there does exist a relationship between "form of inquiry" and "paradigm of activity." (Dr. Hughes is Assistant Director of the Atlantic Institute of Education, Halifax.)*

The problem of how best to organize knowledge for the purposes of instruction is one which has been perennially puzzling for the curriculum scholar. Whether the structure of the curriculum should reflect a separate-subjects organization of knowledge, or the more encompassing "forms" of knowledge, or whether indeed the curriculum should reflect some other manner of organizing knowledge for instruction, are among the questions to be confronted continually. However, while the choices available have clearly been numerous, the school curriculum, particularly at the secondary school level, has traditionally reflected a separate-subjects approach to organizing knowledge for instruction, perhaps more through inertia than through rational selection. And indeed, it would seem that the "disciplines . . . have been reified and institutionalized to such an extent that in curriculum planning we have difficulty thinking beyond them" (Wise, 1966, p. 276).

From time to time, however, there have arisen challenges to the traditional organization, challenges which have resulted in curricula frequently characterized as interdisciplinary or integrated. Such curricula have tended to be varied. They possess in common only the negative attribute of opposition to subject divisions or subject barriers as these are traditionally conceived (Pring, 1973). Their opposition is founded in the belief that the separate-subjects curriculum fosters "a fatal disconnection" (Caswell, 1962, p. 108), that it "fragments a student's world view" (Fraenkel, 1970, p. 376), and that it results in "the creation of a mind torn against



itself" (Wise, 1966, p. 392). These outcomes, it is claimed, are the results of the "artificial barriers" (Salt, 1969, p. 23), "separate compartments" (Taylor, 1970, p. 69), and the "artificially circumscribed packages of information" (Kerr, 1968, p. 26) which characterize the separate-subjects curriculum. Implicit in the argument is the belief that any curricular organization of knowledge which diminishes the strength of the barriers between subjects will inevitably promote the opposite and more desirable effects by permitting the student exposure to what Marashio (1971) has termed the "full dimensional learning experience" (p. 252).

For the most part, however, the assertions represent expressions of faith rather than conclusions based upon evidence. By and large it is ignored that, to the present time, it has not been demonstrated that there is any "a priori theoretical advantage of interdisciplines over disciplines, no matter how we slice them" (Cameron, 1965, pp. 309-310). Nor has it been shown that there is any "empirical evidence to justify any type of discipline curriculum selection and organization — single, correlated or multidisciplinary, broad unified, fused or interdisciplinary — as compared with another" (Rice, 1973, p. 11).

Nevertheless, the literature remains full of the polemics of curriculum organization, all of it based on the assumption that the manner in which knowledge is organized for instruction does, in some measure, constrain the possibilities afforded by the instructional situation, with the resulting learning experiences being different in their major characteristics (Faix, 1964; Montebello, 1964).

The present study was undertaken to subject the assumption to an initial empirical examination. The critical question addressed by the study is: Does there exist a relationship between the manner in which knowledge is organized for instruction and the properties exhibited by the instructional system in which it is situated? Essentially, the task was one of determining whether a student's learning experience is likely to differ in any of its major characteristics as the result of being associated with one type of knowledge organization as opposed to another.

### *Curricular Knowledge Organizations*

It may readily be seen that there are a number of possible ways in which knowledge may be organized so as to help the individual create meaning out of the apparent anarchy of events; the "blooming, buzzing confusion" of Whitehead. There is that organization reflected in the divisions of the Comtian hierarchy and based upon "subject matter" (Comte, 1877), or that based upon Plato's divided line (Cornford, 1967), or Aristotle's four causes. More recently there is that organization found in Tykociner's (1964) zones and areas of knowledge, Hirst's (1965) fields and forms, Phenix's (1964) realms of meaning, and Clark's (1972) general ecology of knowledge.

Recognizing that there are a number of possible ways in which knowledge may be organized, the curriculum builder is faced with the problem of deciding which organization or organizations to select for the purposes of instruction. However, while a wide variety of possibilities for knowledge organization is available to curriculum designers, the choice, in effect, has been historically limited to two, both employing subject matter as



reflected in the Comtian hierarchy as the criterion of categorization. Indeed, the predominance of these two choices has been so evident in curriculum design that it can be suggested that there are two ideal-typical ways in which knowledge may be organized so as to constitute a curriculum. These have been referred to by Bernstein (1971) as educational knowledge codes of the collection and integrated types.

In order to explicate the nature of the codes, Bernstein employs the concept of "classification." Classification, he points out

... does not refer to *what* is classified, but to the relationships between contents. Classification refers to the nature of the differentiation between contents. Where classification is strong, contents are well insulated from each other by strong boundaries. Where classification is weak, there is reduced insulation between contents for the boundaries between contents are weak or blurred. *Classification thus refers to the degree of boundary maintenance between contents.* (p. 49; italics in original)

Where classification is strong, i.e., where contents are clearly bounded and well insulated, standing in a closed relation to each other, the knowledge code is of a *collection* type. Where contents stand in an open relation to each other, with weak or nonexistent boundaries, the knowledge code is of an *integrated* type.

A collection code of the ideal type would appear to exist where a curriculum consists solely of subjects each of which corresponds to a recognized discipline. An integrated code of the ideal type would exist where previously insulated subjects are subordinated to a relational idea which blurs the boundaries between them. Clearly, however, the ideal states will rarely obtain in the existential situation. Nevertheless, it is possible to observe various forms of collection and integration in school classrooms. For example, some measure of integration can be recognized where "knowledge is grouped for pedagogical purposes in four major categories — the natural sciences, the social sciences, mathematics and the humanities" (Bellack, 1965, p. 323). A total curriculum organized in this fashion is rare, but a partial representation is frequently found where the individual social sciences receive consideration as "social studies," the individual identity of each discipline being subordinated to a major relational idea, i.e., the need to explain man's social life.

By comparing the properties of instructional systems (herein referred to as the instructional systemics) employing variously the "social studies" and the separate social science subjects, it was sought to provide an empirical base from which to view the broader problem. The "social studies" and the separate social science subjects were, therefore, viewed as representing a special case of the general phenomenon, the "social studies" representing a measure of integration, and the separate social science subjects a measure of collection.

### *Defining the Instructional System*

The present study was undertaken to investigate the relationship between the curricular knowledge organization (knowledge code) employed in an instructional situation and the characteristics displayed by that situation. From the infinite number of variables in operation, four were selected for particular consideration. These were labelled (a) the



instructional intent, (b) the conditions of “knowing that,” (c) the form of inquiry, and (d) the paradigm of activity. Each was considered as constituting a significant dimension of the instructional process in any classroom, and the social studies or social science classroom in particular.

### *Instructional Intent*

Generally speaking, it may be asserted that a major task of schooling is that of having students acquire knowledge. It is this acquisition of knowledge which comprises the school’s “instructional intent.” Frequently, however, it would appear that schools function according to a rather constrained view of what it is that constitutes knowledge. Often the definition to be inferred from their operations is one of “a rhetoric of conclusions” (Schwab, 1960, p. 185) to be transferred to the student. It is a conception which emphasizes the acquisition of a “compendium of information” (Parker & Rubin, 1966, p. 1) and tends to neglect those cognitive processes which focus upon the creation, utilization, and communication of knowledge.

This distinction, drawn by Ryle (1949) between *knowing that* something is the case and *knowing how* to do something, implies a logical difference between the possession of information and the possession of skills. In the context of the present study the distinction has a particular heuristic value, permitting a ready dichotomization of the notion of instructional intent.

### *Conditions of “Knowing That”*

To claim to know that something is the case may not be simply a question of stating that “such and such is false and so and so is true” (Fen, 1966, p. 164). Nevertheless, many of the knowledge claims that are made, particularly in social studies classes, do appear to adhere to this limited conception (Massialas & Sprague, 1974). A knowledge claim of this type has been termed knowing that in the *weak sense*, and is to be contrasted with knowing that in the *strong sense* (Hintikka, 1962, pp. 18-19) in which case the person claiming to have knowledge must go beyond the initial assertion of what is the case, providing justification for the assertion.

The distinction between knowing that in the *weak sense* and knowing that in the *strong sense* is one employed by Scheffler (1965) who points out that “in the weak sense, knowing that depends solely in having true belief; in the strong sense it requires something further — for example, the ability to back up the belief in a relevant manner, to bring evidence in its support” (p. 9). The conditions of knowing that in the two senses are clearly different. In the *weak sense* there is only the condition of belief; in the *strong sense*, in addition to the condition of belief, there is the condition of evidence, providing what Ayer (1956) calls the right to be sure. Knowing that in the *strong sense*, therefore, is a condition of warranted belief (Conklin, 1967). Conversely, knowing that in the *weak sense* is a condition of unwarranted belief.

### *The Forms of Inquiry*

The importance of knowing as an objective in the educative process has been emphasized with respect to the instructional intent and the conditions of “knowing that.” It is not a notion, however, which may be usefully

employed to describe the day to day activities of teachers and students. Knowing can be viewed as an objective which may serve to guide activities, and when attained it is an achievement fulfilling certain conditions; it is not a kind of mental operation or activity (Ryle, 1949, p. 16 ff.). Day to day instructional activities may be usefully interpreted in terms of the conscious effort to gain knowledge, that is, in terms of the process of inquiry. That is not to say that all knowledge is the result of deliberate inquiry. Clearly, much of what we know we have not deliberately set ourselves to find out (Scheffler, 1965, p. 32), but schooling is a purposive activity in large measure concerned with the attainment of knowledge, and in this context inquiring may be seen to be the "subservient task activity" to knowing.

Morris (1964), in his work on semiotic, has suggested that there are three main types of problems to which inquiry may be directed: (1) problems of what has happened, is happening, or will happen, (2) problems of what to do, and (3) problems of what to accord preferential behaviour. Morris further points out that inquiry will terminate in three different kinds of assertion with respect to these three different kinds of problems. Inquiry into what has happened, is happening or will happen terminates with designative statements and is referred to as *designative inquiry*. Inquiry into what to do terminates with prescriptive statements and is referred to as *prescriptive inquiry*. Inquiry into what to prefer terminates with *appraisive inquiry* (Morris, 1964, p. 26). While conceptually distinct, the forms of inquiry are not separate and, while engaged in any one of them, "one uses the other forms of inquiry at various stages of the process" (Morris, p. 28). The forms of inquiry, therefore, may be regarded as phases of, and emphases within, a more general process of inquiry, considered here as a reflective process involving a transaction between a display and an inquirer, and directed toward solving a problem. (The notion of "display" employed here is derived from Aoki, 1970; Dewey and Bentley, 1949; Jackson, 1966; and Johnson, 1969; and is used to refer to the selected segments of the environment given symbolic representation in the educational knowledge codes.) It is this transaction between student and display which, for the purposes of the present study, is seen as crucial in the instructional situation.

While a distinction may be drawn between prescriptive and appraisive inquiry, it is also possible to regard them as collectively constituting *evaluative inquiry* as opposed to designative inquiry which is *nonevaluative*. It is this latter distinction, i.e., between evaluative and nonevaluative inquiry, which was drawn upon extensively in the study. Again, it should be recognized that these distinctions are emphases within a more general process of inquiry. In evaluative inquiry, the inquirer is concerned with deciding upon a preferential course of behaviour, with identifying the objects and actions which will receive preferential consideration. In order to make his decisions, he will be dependent upon accurate information which can only be acquired through designative inquiry. In this way, nonevaluative or designative inquiry can be viewed as a phase, albeit a prerequisite phase, in the process of evaluative inquiry.



### *Paradigm of Activity*

It has been noted in the previous section that the fundamental interaction in the process of instruction may be viewed as taking place not among people but between the individual and the “display.” However, it was further noted that the instructional transaction is mediated by the “instructor” who may be personally present or absent from the instructional situation, i.e., he may be a teacher actively engaged in instructing children in a classroom situation, or he may be the “author” of materials employed by students in the process of inquiry. In both cases, the instructor, with the student, builds a relationship through which the transaction is mediated. It is the character of this relationship which is considered under the heading of “paradigm of activity.” As used here, the idea of a “paradigm” is descriptive rather than prescriptive, and refers to the major characteristics exhibited by a given phenomenon, the phenomenon in this case being the relationship through which the classroom transaction is mediated. In the restricted sense of the study, a paradigm of activity was considered to be that dominant pattern of interaction, or way of proceeding, which is reflected in the authority relationships existing between the display and/or teacher, and the student.

The importance of authority relationships, particularly those existing between teachers and students, have been emphasized in attempts to describe how classrooms operate (Dreeben, 1972, pp. 457-463). The relationship has been variously characterized along continua, from democratic to autocratic, integrative to dominative, student-centered to teacher-centered. However, the tradition of democratic-authoritarian studies, stemming largely from the work of Lewin (1943) and Lippett and White (1943), has been inconclusive in terms of any measure of classroom effectiveness, especially when the latter has been approached in terms of “productivity” (Anderson, 1959).

While the tradition of research referred to above has focused upon the ideas of a “style of leadership” within a broad context, the conceptions of teacher-regulated and group-regulated classrooms derived from it were of particular value in the present study, permitting a dichotomous consideration of the paradigm of activity. This was particularly important in view of the fact that one dimension of the dichotomy, the teacher-centered, has tended to be associated with an educational knowledge code of the collection variety and the other, the group-regulated, with an integrated code.

### *Method*

The research design provided for five separate and independently conducted naturalistic and quantitative case studies. In each of the cases the relationship obtaining between the curricular knowledge organizations employed and the other elements of the abstracted instructional system was examined, and invariances across cases were mapped. It was hypothesized that there exists a relationship between the knowledge code employed in an instructional situation and the relative emphasis given to: (i) *knowing how* and *knowing that*, (ii) “knowing that” in the *weak sense* and “knowing that” in the *strong sense*, (iii) *evaluative* inquiry and *nonevaluative* inquiry, (iv) *teacher-centered* activity and *group-regulated* activity.

### *Sample*

Each case was composed of two intact classes, one categorized as employing a collection knowledge organization and the other categorized as employing an integrated knowledge organization. The total sample was comprised of ten intact classes selected through purposive sampling, i.e., classes were selected because they exhibited the attributes relevant to the study.

The use of an equal probability of selection method with such a small sample would not have provided any useful measure of control over teacher style which was viewed as a potentially confounding variable. Through purposive sampling it was possible to identify those teachers who were employing an integrated knowledge organization in one classroom situation and a collection knowledge organization in another classroom at the same grade level. It was hoped that the selection of two classes taught by the same teacher would provide a measure of control over variation in “teacher style” within a single case. An attempt to involve precisely the same students in the two classes comprising each case was aborted when no such instances were found to exist, and only intact classes could be selected. Of the classes which ultimately composed the sample the degree of overlap in student participation in the two situations ranged from 0 to more than 60 per cent.

An initial pool of 14 teachers in six school districts in central Alberta was identified. The teachers comprising the initial pool were those involved in teaching both social studies and one of the social science options (geography, psychology, or sociology) at the grade eleven level. It was thought that the social studies courses might prove to be a source of classes employing integrated knowledge organizations, and social science courses a source of classes employing collection knowledge organizations. An analysis of the courses led to five teachers being tentatively identified as employing an integrated knowledge organization in their social studies classes and a collection knowledge organization in their social science classes.

The basis of discrimination between the integrated and collection knowledge organizations was linked to the concept of “classification” which refers to the strength of boundaries between contents. Boundaries are designated weak when a number of perspectives (psychological, sociological, economic, etc.) are brought to bear on a single topic or problem being considered by a class, and thereby constitute an integrated knowledge organization. Boundaries are designated strong when a topic or problem is considered from a single perspective (i.e., psychological or sociological or economic, etc.) and result in a collection knowledge organization.

Classes were categorized following an examination of textual materials (textbooks, handouts, resource units, assignments and examinations) and observations of the classrooms in action. However, the tentative nature of the initial sampling must be stressed. It was a concern at this point in the development of the study that teacher behaviour with respect to the use of integrated and collection knowledge organizations might be extremely variable over time, though this proved not to be the case. A strict monitoring of classes in terms of content “classification” was, therefore, conducted during the period of data collection.



The ultimate composition of the five cases involved 10 classes, five of which were grade eleven social studies and five of which were social science options. The five option classes were two sociology, two geography, and one psychology. All the teachers were male, having a range of experience from one to twelve years, and of academic qualifications from three years of university to a master's degree. Two of the case studies were conducted in a suburban high school, two in small town high schools, and one in a small rural school. In total the study involved 186 students ranging in age from 14.8 to 18.7 years from a wide range of socioeconomic backgrounds, with approximately equal numbers of males and females.

### *Instrumentation*

Data were collected by five trained observers employing the *System for the Analysis of Classroom Transactions*<sup>1</sup>, a multidimensional category system designed to permit the quantitative description of classroom transactions in terms of the variables relevant to the present study by means of direct observation.

The SACT was developed in order to permit a description of classroom behaviour in terms of instructional intent, conditions of "knowing that," form of inquiry, and paradigm of activity, each of which is represented in the system in terms of dichotomous categories.

For purposes of coding, the units of analysis of the SACT are "analytic" rather than "phenomenal" (Biddle, 1967) and are referred to as *ventures*. The term is used to delineate a unit of classroom interaction which focuses upon a single notion or subject or topic. When there is a shift away from one focus of attention to another, (from one topic to another, or from one aspect of a topic to another) a change of ventures is deemed to have taken place. Typically, a social studies or social science class of approximately 80 minutes can be expected to produce between 15 and 30 ventures.

As indicated, each venture may be classified along each of the four dimensions of the system. For example, in a venture which consists of

TABLE 1  
DIMENSIONS AND CATEGORIES OF THE SACT

Dimensions	Categories
Instructional Intent	Knowing How Knowing That
Conditions of "Knowing That"	Strong Weak
Form of Inquiry	Evaluative Nonevaluative
Paradigm of Activity	Teacher-Centered Group-Regulated

student A asking student B to justify his assertion that Canada should prohibit further immigration by people of nonwhite racial origin, and receiving an appropriate response, the pattern of classification would be: (i) knowing that, (ii) strong sense, (iii) evaluative inquiry, and (iv) group-regulated. However, since the SACT is a category system, observers are frequently called upon to make judgements requiring high levels of inference. In order to increase the validity and reliability of such high inference judgements, a number of “significant indicators” (in effect, low level signs) for each category have been identified to provide guidance for observers.<sup>2</sup>

Five persons, one of whom was the researcher, were used as observers. Before data collection the observers engaged in approximately 15 hours of training using the SACT. With observers having no training whatever and being provided with only the names of the categories and a single statement description for each dimension of the system, the coefficients of inter-observer reliability<sup>3</sup> ranged from 0.56 to 0.81 with a median of 0.66. After training, using transcripts, and audio and video recordings of classroom interaction, the coefficients ranged from 0.93 to 1.00 with a median of 0.97. At the end of the 50-day period of data collection, coefficients of inter-observer reliability ranged from 0.76 to 0.97 with a median of 0.81.

Intra-observer reliability was consistently high during both the training and data collection periods. Intra-observer reliability calculated during a ten-day interval at the end of training yielded coefficients ranging from 0.93 to 1.00 with a median of 0.98. After the 50-day period of classroom observation, coefficients ranged from 0.93 to 0.98 with a median of 0.95.

### *Procedures*

Data in each of the cases were collected by one of the observers. Each of the two classes comprising each case study was observed and coded for five periods each of eighty minutes in length. Periods were not randomly selected. Observers, in conjunction with the teachers, selected those periods likely to provide large measures of teacher-student and student-student interaction, thereby excluding those learning activities such as field trips and independent study which are not amenable to analysis using the SACT.

### *Results and Discussion*

The unit of analysis referred to is that unit of classroom discourse referred to in the study as a “venture.”

### *Instructional Intent*

In none of the five case studies was there found to be a systematic variation between the relative emphasis placed upon instructional intent of the “knowing how” and “knowing that” types. In both types of classes there was almost exclusive attention given to instructional intent of the “knowing that” type. Of 1,153 ventures analyzed in the five case studies, only 24 (approximately 2 per cent) were classified as “knowing how,” and would seem to confirm Soltis’s (1968) and Holloway’s (1975) assertions that in spite of the considerable concern for process (“knowing how”) in the



TABLE 2  
CHI-SQUARE TEST OF INDEPENDENCE BETWEEN KNOWLEDGE  
ORGANIZATION AND INSTRUCTIONAL INTENT

Knowledge Organization	Total Number of Ventures	Instructional Intent		$\chi^2$
		Knowing How	Knowing That	
<u>Case Study I</u>				
Integrated	109	0	109	
Collection	94	0	94	--
<u>Case Study II</u>				
Integrated	108	5	103	
Collection	116	0	116	3.58
<u>Case Study III</u>				
Integrated	133	10	123	
Collection	130	6	124	0.53
<u>Case Study IV</u>				
Integrated	149	1	148	
Collection	73	2	71	0.40
<u>Case Study V</u>				
Integrated	124	0	124	
Collection	117	0	117	--

Critical value of  $\chi^2$  (df = 1) at  $p < 0.05 = 3.84$ .

literature, "knowing that" is *the* basic type whether the mode of curricular knowledge organization is collection or integrated.

#### *The Conditions of "Knowing That"*

In three of the case studies (I, II, and III) there was found to be no difference between the integrated and collection classes with respect to the emphasis placed upon "knowing that" in its strong and weak senses. In cases IV and V, differences were observed ( $p < .01$  and  $< .001$  respectively). In both of these latter instances it was found that ventures categorized as "knowing that" in the *weak* sense occurred to a significantly greater degree in the integrated classes, although the proportion of weak to the total number of ventures was less than in either case study I or II, in which no difference between collection and integrated classes were observed.

The lack of consistency of findings among cases suggests that there is no systematic variation in the relative emphasis given "knowing that" in its *strong* and *weak* senses between classes employing collection and integrated knowledge codes.

#### *The Form of Inquiry*

In four of the case studies it was found that a significantly greater emphasis was placed upon evaluative inquiry in the integrated classes than

TABLE 3  
CHI-SQUARE TEST OF INDEPENDENCE BETWEEN KNOWLEDGE  
ORGANIZATION AND THE CONDITIONS OF KNOWING THAT

Knowledge Organization	Total Number of Ventures	Conditions of Knowing That		$\chi^2$
		Strong Sense	Weak Sense	
<u>Case Study I</u>				
Integrated	109	74	35	0.50
Collection	94	69	25	
<u>Case Study II</u>				
Integrated	103	103	0	1.95
Collection	116	112	4	
<u>Case Study III</u>				
Integrated	123	90	33	0.87
Collection	124	98	26	
<u>Case Study IV</u>				
Integrated	148	130	18	7.87*
Collection	71	71	0	
<u>Case Study V</u>				
Integrated	124	104	20	18.51**
Collection	117	117	0	

\*  $p < .01$

\*\*  $p < .001$

was the case in collection classes. The findings would seem to lend some measure of support to the hypothesis that there exists a relationship between the curricular knowledge organization employed and the relative emphasis given to evaluative and nonevaluative inquiry. However, the finding is somewhat confounded by the fact that the social studies curriculum of the Province of Alberta is explicitly “valuing-oriented” and the emphasis on valuative inquiry in the integrated classes (all of which were social studies) may be related more to this curricular orientation than to the curricular knowledge organization employed.

*The Paradigm of Activity*

In all five case studies significant differences were observed between the collection and integrated classes in terms of the numbers of ventures classified as *teacher-centered* or *group-regulated*. In four cases it was found that *group-regulated* ventures occurred to a significantly greater extent in the integrated classes. In a single case (case study II), it was found that group-regulated ventures occurred to a significantly greater degree in the collection class.

The high proportion of teacher-centered to group-regulated ventures in the collection classes tends to confirm the findings of Bellack (1966) and Hughes (1968) that teachers make little use of student ideas and that the



TABLE 4  
CHI-SQUARE TEST OF INDEPENDENCE BETWEEN KNOWLEDGE  
ORGANIZATION AND FORM OF INQUIRY

Knowledge Organization	Total Number of Ventures	Form of Inquiry		$\chi^2$
		Evaluative	Nonevaluative	
<u>Case Study I</u>				
Integrated	109	36	73	14.76**
Collection	94	9	85	
<u>Case Study II</u>				
Integrated	108	8	100	1.3
Collection	116	15	101	
<u>Case Study III</u>				
Integrated	133	32	101	3.82*
Collection	130	18	112	
<u>Case Study IV</u>				
Integrated	149	13	136	5.27*
Collection	73	0	73	
<u>Case Study V</u>				
Integrated	124	16	108	14.2*
Collection	117	0	117	

\*  $p < .05$

\*\*  $p < .001$

fundamental paradigm of classroom activity is one of teacher initiation and student response. On the other hand, the findings in the integrated classes suggest that while the teacher clearly exerts a controlling influence over instructional activities, an important proportion of instructional activities is, in fact, regulated by the group. The evidence from the cases studied might be taken to indicate that the collection classes, based on clearly bounded disciplines, are more constraining than integrated classes in terms of the possibilities afforded in the instructional situation.

#### *Summary and Conclusions*

The focus of this investigation derives from a frequently expressed concern in curriculum discourse that the traditional separate-subjects approach to organizing knowledge for instruction is no longer adequate. The argument usually advanced by opponents of the separate-subjects approach is that such subjects are fragmentary and compartmentalized; that there is a need to proceed beyond that curriculum design characterized as "a process of layering society's new knowledge and arranging for feeding it." Further, they would claim that together with other societal tendencies toward specialization, they have resulted in a "depersonalization and demoralization, disintegration, and disorientation, the brink of world holocaust and the brink of ecological suicide, the alienation of



TABLE 5  
CHI-SQUARE TEST OF INDEPENDENCE BETWEEN KNOWLEDGE  
ORGANIZATION AND THE PARADIGM OF ACTIVITY

Knowledge Organization	Total Number of Ventures	Paradigm of Activity		$\chi^2$
		Teacher- Centered	Group- Regulated	
<u>Case Study I</u>				
Integrated	109	43	66	44.4***
Collection	94	81	13	
<u>Case Study II</u>				
Integrated	108	69	39	8.1*
Collection	116	51	65	
<u>Case Study III</u>				
Integrated	133	87	46	4.91*
Collection	130	102	28	
<u>Case Study IV</u>				
Integrated	149	54	95	28.43***
Collection	73	55	18	
<u>Case Study V</u>				
Integrated	124	79	45	19.52***
Collection	117	104	13	
* $p < .05$				
** $p < .01$				
*** $p < .001$				

instructionalized man and the deprivation of the human spirit” (Haas, 1975, p. 5).

The present study sought to determine whether the process of education in classes using collection (separate-subject) knowledge organizations was likely to differ in any of four dimensions from classes using integrated knowledge organizations. The results clearly indicate that there is no difference between collection and integrated classes in terms of “instructional intent” or “the conditions of knowing that.” The results concerning the “form of inquiry” and the “paradigm of activity” suggest that evaluative inquiry and group-regulated activity are likely to receive greater emphasis within the integrated classes. However, the results of the case studies, viewed collectively, would seem to indicate that any assertion claiming the mode of knowledge organization as a variable controlling and shaping the nature of instructional systems in some inexorable fashion is clearly untenable. At best, it is one of a number of variables shaping the situation.

The debate concerning how knowledge ought to be organized for the purposes of instruction is inevitably linked to conceptions of how and what people ought to be. Basically, there appear to be three assumptions. The first is that integration is a “good” thing, whether in the sense of integrated



days, integrated subjects, racial integration, an integrated personality, or even integrated circuits. Consequently, people ought to be “integrated” quite simply because the idea connotes something positive, as opposed to its opposite, segregation or possibly disintegration, which connotes something negative. The second assumption is that the knowledge organization with which any individual transacts is a contributory factor in the process of becoming “integrated.” Finally, it is assumed that knowledge organizations, which are in themselves more integrated, contribute to the integration of individuals. It might well be, however, that knowledge organizations that are already integrated rob the inquirer of the experience of integrating and place him in the role of the passive consumer. Knowledge organizations that are not already integrated may force the student into dynamic inquiry, seeking to create meaning out of the “great, blooming, buzzing confusion.”

#### Notes

1. Copies of the *SACT* and *Guidelines for Observers* may be obtained directly from the author.
2. See *System for Analyzing Classroom Transactions: Guidelines for Observers*.
3. The coefficient of inter-observer reliability was arrived at using procedures advocated by Smith, Meux, Coombs, and Nuthall, in “A Tentative Report on the Strategies of Teaching,” Bureau of Educational Research, College of Education, University of Illinois, Urbana, n.d., p. 41.

$$C = \frac{A_1 + A_2}{B_1 + B_2}$$

C = Coefficient of agreement

A<sub>1</sub> = The number of ventures assigned to a particular category by observer #1 which are also assigned the same category by judge #2.

A<sub>2</sub> = The number of ventures assigned to a particular category by observer #2 which are also assigned the same category by judge #1.

B<sub>1</sub> = The total number of ventures identified by observer #1

B<sub>2</sub> = The total number of ventures identified by observer #2

#### References

- Anderson, R. C. Learning in discussions: A resume of authoritarian-democratic studies. *Harvard Educational Review*, 1959, 29, 201-215.
- Aoki, T. A curriculum and instructional design. Paper presented at Project Canada West Workshop, Edmonton, 1970.
- Ayer, A. J. *The problem of knowledge*. Harmondsworth, England: Penguin, 1956.
- Bellack, A. What knowledge is of most worth? *The High School Journal*, 1965, 48, 318-332.
- Bellack, A., Kleibard, H. M., Hyman, R. T., & Smith, F. L. *The language of the classroom*. New York: Teachers College Press, 1966.
- Bernstein, B. On the classification and framing of educational knowledge. In M. F. D. Young (Ed.), *Knowledge and control: New directions for the sociology of education*. London: Collier-MacMillan, 1971.
- Biddle, B. J. Methods and concepts in classroom research. *Review of Educational Research*, 1967, 37, 337-357.
- Cameron, W. B. The hydra-headed curriculum: The proliferation of inter-disciplinary studies. *Journal of Higher Education*, 1965, 36, 307-312.
- Caswell, H. L. Difficulties in defining the structure of the curriculum. In A. H. Passow (Ed.), *Curriculum crossroads*. New York: Teachers College Columbia, 1962.
- Clark, J. W. The general ecology of knowledge in curriculums of the future. In E. Laszlo (Ed.), *The relevance of general systems theory*. New York: George Brazillier, 1972.
- Comte, A. *Cours de philosophie positive*. Paris: Bailliere, et fils, 1877.

- Conklin, K. R. The integration of the disciplines. *Educational Theory*, 1967, 16, 225-238.
- Cornford, F. MacD. *The Republic of Plato*. London: Oxford University Press, 1967.
- Dewey, J., & Bentley, A. F. *Knowing and the known*. Boston: Beacon Press, 1949.
- Dreeben, R. The school as a workplace. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Faix, T. L. Toward a science of curriculum: Structural-functional analysis as a conceptual system for theory and research (Doctoral dissertation, University of Wisconsin, 1964). Ann Arbor, MI: University Microfilms, 1964, No. 64-10, 230.
- Fen, S. N. "Knowing that" rediscovered and its place in pedagogy reassigned. *Educational Theory*, 1966, 16, 163-168.
- Fraenkel, J. R. Program definition: Logic and process. *The High School Journal*, 1970, 53, 375-384.
- Haas, J. D. For lack of a loom: Problems in integrating knowledge. *School Science and Mathematics*, 1975, 75 (1), 4-14.
- Hintikka, J. *Knowledge and belief*. Ithaca, New York: Cornell University Press, 1962.
- Hirst, P. H. Liberal education and the nature of knowledge. In R. D. Archambault (Ed.), *Philosophical analysis and education*. London: Routledge and Kegan Paul, 1965.
- Holloway, O. *Problem-solving: Toward a more humanizing curriculum*. Philadelphia: Franklin Publishing, 1975.
- Hughes, M. M. What is teaching? One viewpoint. In R. J. Hyman (Ed.), *Teaching vantage points for study*. Toronto: J. B. Lippincott, 1968.
- Jackson, P. W. The way teaching is. In P. W. Jackson (Ed.), *The way teaching is*. Washington: ASCD, 1966.
- Johnson, M. The translation of curriculum into instruction. *Journal of Curriculum Studies*, May 1969, 1, 115-131.
- Kerr, J. F. *Changing the curriculum*. London: University of London Press, 1968.
- Lewin, K. Psychology and the process of group living. *Journal of Social Psychology*, 1943, 17, 113-131.
- Lippett, R., & White, R. K. The "social climate" of children's groups. In R. G. Barker, J. S. Kounin & H. F. Wright (Eds.), *Child behaviour and development*. New York: McGraw-Hill, 1943.
- Marashio, P. The humanities today: Moving toward a new direction. *The Clearing House*, 1971, 46, 252-253.
- Massialas, B. B., & Sprague, N. F. Teaching social issues as inquiry: A clarification. *Social Education*, 1974, 38, 10-19; 34-35.
- McClure, R. M. The reforms of the fifties and sixties: A historical look at the war past. In R. M. McClure, (Ed.), *The curriculum: Retrospect and prospect*. Chicago: N.S.S.E., Seventieth yearbook, 1971.
- Montebello, M. S. The thinking and practice of Ohio city elementary school principals in reference to unit teaching and the integrative curriculum organization (Doctoral dissertation, The Ohio State University). Ann Arbor, MI: University Microfilms, 1964, No. 65-3892.
- Morris, C. *Signification and significance: A study of the relation of signs and values*. Cambridge, Mass.: M.I.T. Press, 1964.
- Parker, J. C., & Rubin, L. J. *Process as content: Curriculum design and the application of knowledge*. Chicago: Rand McNally, 1966.
- Phenix, P. H. *Realms of meaning: A philosophy of the curriculum for general education*. New York: McGraw-Hill, 1964.
- Pring, R. Curriculum integration: The need for clarification. *The New Era*, 1973, 54, 59-64.
- Rice, M. J. The disciplines-languages approach to curriculum development. Paper presented at the Curriculum Symposium of the American Education Research Association, New Orleans, February, 1973.



- Ryle, G. *The concept of mind*. London: Hutchison's University Library, 1949.
- Salt, J. Problems of integrated education. *Trends in Education*, 1969, 16, 23-27.
- Scheffler, I. *Conditions of knowledge: An introduction to epistemology and education*. Chicago: Scott Foresman and Co., 1965.
- Schwab, J. J. Inquiry, the science teacher, and the educator. *School Review*, 1960, 68, 176-195.
- Smith, B. O., Meux, M. O., Coombs, J., & Nuthall, G. A tentative report on the strategies of teaching. Urbana, Illinois: Bureau of Educational Research, College of Education, University of Illinois, n.d.
- Soltis, J. F. *An introduction to the analysis of educational concepts*. Reading, Mass.: Addison-Wesley, 1968.
- Taylor, H. Inside Buckminster Fuller's universe. *Saturday Review*, 1970, 53 (May 2), 56-57; 69-70.
- Tykociner, J. T. Zetetics and areas of knowledge. In S. Elam (Ed.), *Education and the structure of knowledge*. Chicago: Rand McNally, 1964.
- Wise, G. Integrative education for a dis-integrated world. *Teacher's College Record*, 1966, 67, 391-401.

JEAN H. YOUNG

*The University of Alberta*

## The Curriculum Decision-Making Preferences of School Personnel

*The decentralization of curriculum decision-making, which has been accepted in principle in many provinces, depends upon the participation of school personnel at local levels. This article explores the probable extent of participation of school personnel in curriculum decision-making by presenting data regarding: (a) preferences for seven kinds of curriculum work, (b) preferences for five levels of curriculum decision-making, and (c) perceptions of the attractiveness of curriculum work. The data were collected by means of a measuring instrument which was distributed to personnel in 174 public and separate schools in Alberta. The 1,268 respondents and the 139 schools in which they worked were found to be representative of school personnel and schools throughout the province, thereby increasing the generalizability of the findings. Distinct preferences for the seven kinds of curriculum work and the five levels of curriculum decision-making are identified in the article as well as the prevailing attitude toward participation. Implications of the findings for the probable extent of participation of school personnel in curriculum decision-making are discussed. (Dr. Young is Associate Professor in the Department of Elementary Education at The University of Alberta.)*

In recent years, a trend has developed across Canada to decentralize curriculum decision-making. Driving forces behind the trend are the increasing professionalism of teachers, the desire of stake-holders in education to participate in decision-making which affects their schools and children, and the positive outcomes of well-publicized local curriculum projects.

Using the province of Alberta as an example, evidence of these forces can be found in policy statements of The Alberta Teachers' Association (1976), recommendations resulting from a major conference of stakeholders in Alberta education ("Curriculum Decision-Making in Alberta," 1974), and the local curriculum work of Project Canada West (Miller & Dhand, 1973).

The trend to decentralize curriculum decision-making was initiated by educational policy makers in Alberta and other provinces. However, the



long term success of the trend depends not on the policy makers, but on the thousands of school personnel whose participation at local levels is a basic requirement of decentralization. As Sullivan points out: "Decentralization of curriculum development has as a major tenet the systematic increase in decision-making responsibilities for teachers and principals" (1975, p. 15).

However, the orientation of school personnel appears to be toward instructional rather than curricular concerns. For teachers in particular, the working day revolves around the exigencies of classroom life. Morrison, Osborne, and McDonald (1977) point out that, as a result, a teacher's informational field regarding curriculum matters is quite restricted. Bounded by the four walls of his classroom, the teacher maintains few external contacts with professionals in his own or related fields. Nor does he have time to visit new programs or projects.

It is not surprising, then, that teachers in a study by Olson and Kitto (1977) gave lip service to the importance of curriculum development but devoted very little time to it. Furthermore, there is evidence that additional time to participate in curriculum decision-making would not necessarily counteract the strong orientation of teachers toward instruction. For example, when Lortie (1975) asked teachers what they would do with a gift of additional work time, he found that they would spend it on activities directly related to their individual classrooms rather than to broader curriculum development work.

On a more positive note, teachers in a recent study by Ponder and Bulcock (1976) expressed a desire for greater participation in curriculum decision-making. It is interesting to note, however, that the greatest discrepancy between actual and desired participation occurred with respect to determining texts and instructional material — a kind of curriculum work which has immediate implications for classroom activities. The discrepancy was considerably smaller for determining the basic outline and detailed content of a curriculum.

To what extent, then, can school personnel be expected to participate in curriculum decision-making? Data drawn from a recent study of the curriculum decision-making preferences of Alberta school personnel approach this question from three different perspectives (Young, 1977).

First, the kinds of curriculum work that are attractive to school personnel are identified. The term "curriculum decision-making" is a very broad concept which masks the possibility that there are different kinds of curriculum work, each one requiring a specific set of decisions. It is possible that school personnel are more willing to participate in certain kinds of curriculum work than in others. Findings of this nature have implications for the scope of participation of school personnel in curriculum decision-making.

Second, the levels at which school personnel wish to perform the various kinds of curriculum work are identified. Curriculum decision-making can be decentralized to different levels: regions of a province, school districts, schools, and individual classrooms. School personnel may welcome responsibility for particular kinds of curriculum work in their schools, but believe that other kinds of curriculum work would be more appropriately performed at other levels. In that case, the extent of participation of school

personnel will vary according to the kind of curriculum work that is proposed.

Third, the attitudes of school personnel toward the opportunity to become more active in curriculum work are identified. Some school personnel may welcome the opportunity, while others react in a “Well, if I *have* to” manner. Still others may refuse to become more active in curriculum work.

Specifically, then, this article explores the probable extent of participation of school personnel in curriculum decision-making by identifying:

1. Preferences for seven kinds of curriculum work.
2. Preferences for five levels of curriculum decision-making.
3. Perceptions of the attractiveness of curriculum work.

The study from which the data were drawn was intended to be exploratory in nature and to provide information which educators with a stake in decentralized curriculum decision-making may use to direct the course of decentralization in their provinces.

### *Method*

#### *Conceptualization of Curriculum Decision-Making*

The problem underlying this study was the prediction of the probable extent of participation of school personnel in curriculum decision-making. In order to solve that problem, it was necessary to operationalize the ambiguous term “curriculum decision-making” for research purposes.

The researcher, therefore, identified seven kinds of curriculum work around which the many curriculum decisions are clustered. Two criteria were used to determine inclusion of a particular kind of curriculum work.

The first criterion was consistency with the definition of curriculum as *a plan outlining the objectives and content of a subject which is available to learners in school*. This definition adheres to the viewpoint that there is a difference between curriculum and instruction. That is, curriculum decision-making takes place before a teacher interacts with students in his classroom (Beauchamp, 1975; Johnson, 1969).

The second criterion was potential performance of the curriculum work at any level of decision-making within a province: *province, region, school district, school, and classroom*. On a continuum ranging from highly centralized (province) to completely decentralized (classroom), each level represents an additional step in the decentralization process.

Seven kinds of curriculum work met the criteria. Five kinds pertain to individual subjects — creating a curriculum, selecting a curriculum, adapting a curriculum, translating a curriculum into instruction, and winning support for a curriculum. A sixth kind of curriculum work — organizing a total program — is concerned with the relationships among individual subjects. The final kind of curriculum work — evaluating curriculum decision-making — applies either to an individual subject or to the total program of which each subject is a part.

The validity of the conceptualization was supported by the fact that school personnel who participated in pilot testing the measuring instrument



based on the conceptualization had no difficulty recognizing the different kinds of curriculum work and distinguishing between them.

### *The Measuring Instrument*

The conceptualization served as the basis for construction of a measuring instrument. The format was that of a simple brochure which could be self-administered in ten minutes or less. The contents were tested and revised during three stages of pilot testing with Alberta school personnel.

The cover page of the brochure advertised the possibility of "Workshops on Curriculum Decision-Making." One side of the brochure contained brief descriptions of seven workshops corresponding to the seven kinds of curriculum work. The other side of the brochure was a questionnaire in which the respondents listed the three workshops they would be most likely to attend, identified the level of curriculum decision-making they would like each workshop they chose to concentrate on, and indicated whether the opportunity to become more active in curriculum work was "very attractive," "somewhat attractive," or "not attractive." The remainder of the questionnaire consisted of multiple choice questions regarding the respondents' school situations, types of employment, and professional backgrounds.

The workshop format was chosen in order to avoid the frequently glib responses associated with checklists and to give a more realistic setting for data collection. Although the respondents made no actual commitment to attend a workshop, the possibility of future commitment existed. This possibility was enhanced by the fact that the measuring instrument was sent to respondents under the auspices of The Alberta Teachers' Association. There is also some evidence that teachers regard workshops favourably. For example, teachers in the Olson and Kitto study (1977) perceived workshops as having the second most potential (after "other teachers") for helping them with curriculum work.

### *The Study Sample*

Theoretically, a random sample of school personnel throughout the province would lead to the greatest generalizability of the findings. However, discussion during pilot testing indicated that the return from a direct mailing would be low. Therefore, a random sample was drawn from the public and separate schools of the province. A total of 174 schools, or 15% of the population, was selected. The measuring instrument was then distributed in each school to full-time personnel with valid Alberta teaching certificates by The Alberta Teachers' Association representative in that school.

Administrators, librarians, and guidance counselors, as well as teachers, received the measuring instrument. Thus, the term "school personnel" refers to educators who fill a variety of positions in the individual schools of a district.

A total of 139 out of 174 randomly-selected schools participated in the study. These schools represented 80% of the study sample. A total of 1,292 individuals completed the measuring instrument, representing

approximately 50% of the personnel in participating schools. Fully 98% of the questionnaires (or 1,268) were usable, attesting to the quality of the pilot testing procedure.

If generalizations were to be drawn regarding the attitudes of school personnel toward curriculum work, it was important that the study respondents be representative of school personnel throughout the province. Therefore, the respondents' characteristics were compared with the characteristics of school personnel throughout Alberta as indicated by data supplied by the Alberta Department of Education (1976).

The schools in which the respondents worked were found to be representative with respect to type of school, size of community, and size of school. The individuals who participated in the study were representative of Alberta school personnel with respect to role, years of employment in schools, years of postsecondary education, and sex. The overall conclusion was drawn that the respondents may be considered representative of the general population of Alberta school personnel.

Results

This section identifies the preferences of school personnel for seven kinds of curriculum work and five levels of curriculum decision-making as well as perceptions of the attractiveness of curriculum work. Since 79.9% of the respondents were classroom teachers, the findings are dominated by their point of view. The responses of administrators, librarians, and guidance counselors differed to some extent from those of classroom teachers. However, only the aggregate responses are reported in this section.

Preferences for Seven Kinds of Curriculum Work

The relative popularity of the seven kinds of curriculum work, as indicated by choices of workshop, is shown in Table 1. The table records the number of times each workshop was chosen by combining first, second, and third choices. Both absolute numbers and percentages are given. The total

TABLE 1  
DISTRIBUTION OF PREFERENCES FOR SEVEN KINDS  
OF CURRICULUM WORK<sup>a</sup>

Preference Rank	Kind of Curriculum Work	Absolute Frequency	Percentage of Total
First	Translating a curriculum into instruction	863	24.1
Second	Creating a curriculum	798	22.3
Third	Adapting a curriculum	588	16.5
Fourth	Organizing a total program	530	14.8
Fifth	Selecting a curriculum	400	11.2
Sixth	Evaluating curriculum decision-making	237	6.6
Seventh	Winning support for a curriculum	158	4.4
Total		3574	100.0

<sup>a</sup> Totals of first, second, and third choices



does not equal three times the number of respondents (3 x 1,268) because some respondents chose only one or two workshops rather than three.

The three choices of workshop were given equal weighting. This was consistent with the purpose of the study which was to make broad distinctions among the preferences of school personnel.

Distinct preferences for each of the seven kinds of curriculum work can be identified. These preferences clustered into three groups indicating greatest, least, and an intermediate degree of interest.

Preferences for Five Levels of Curriculum Decision-Making

After the respondents had selected the three workshops they would be most likely to attend, they indicated the levels of decision-making on which they would like the workshops they chose to concentrate. In effect, the respondents were indicating the degree of decentralization they felt was desirable for the seven kinds of curriculum work.

Analysis of the data proceeded in two stages. First, preferences for the five levels of curriculum decision-making were studied without regard for specific kinds of curriculum work. Table 2 records in absolute numbers and percentages the number of times each level was chosen. Again, due to omitted responses, the total does not equal three times the number of respondents.

TABLE 2  
DISTRIBUTION OF PREFERENCES FOR FIVE LEVELS  
OF CURRICULUM DECISION-MAKING <sup>a</sup>

Preference Rank	Level of Curriculum Decision-Making	Absolute Frequency	Percentage of Total
First	School district	923	28.3
Second	Province	752	23.1
Third	School	684	21.0
Fourth	Classroom	469	14.4
Fifth	Region	434	13.3
Total		3262	100.0

<sup>a</sup> Totals of levels selected for all workshop choices

In the second stage of analysis, the level of curriculum decision-making preferred for each kind of curriculum work was identified. Table 3 records in absolute numbers and percentages the number of times each level was chosen for each kind of curriculum work.

This stage of analysis clarified the overall preferences reported in Table 2. It indicated, for example, that the overall preference for the *school district* extends over six different kinds of curriculum work. On the other hand, the overall preference for the *province* as second choice focuses primarily on only two kinds of curriculum work. Curiously, preference for the *school* as overall third choice was not clarified by this stage of analysis. That is,

TABLE 3  
LEVELS OF DECISION-MAKING SELECTED FOR THE SEVEN KINDS  
OF CURRICULUM WORK<sup>a</sup>

Kinds of Curriculum Work	Province		Region		School District		School		Classroom		Totals	
	Abs. No.	% of Row Total	Abs. No.	% of Row Total	Abs. No.	% of Row Total	Abs. No.	% of Row Total	Abs. No.	% of Row Total	Abs. No.	% of Grand Total
Translating a curriculum into instruction	110	14.0	83	10.6	208	26.5	179	22.8	205	26.1	785	24.3
Creating a curriculum	193	26.4	115	15.8	184	25.2	147	20.1	91	12.5	730	22.6
Adapting a curriculum	66	12.6	61	11.6	199	37.9	131	25.0	68	13.0	525	16.2
Organizing a total program	193	39.7	57	11.7	108	22.2	85	17.5	43	8.8	486	15.0
Selecting a curriculum	78	21.8	68	19.0	104	29.1	67	18.8	40	11.2	357	11.0
Evaluating curriculum decision-making	54	25.1	35	16.3	70	32.6	43	20.0	13	6.0	215	6.6
Winning support for a curriculum	35	25.7	14	10.3	49	36.0	29	21.3	9	6.6	136	4.2
Totals	729	22.5	433	13.4	922	28.5	681	21.1	469	14.5	3234	100.0

<sup>a</sup> Totals of all workshop and level choices

respondents apparently did not consider the school as the optimum level for any particular kind of curriculum work.

*Perceptions of the Attractiveness of Curriculum Work*

After indicating their preferences for seven kinds of curriculum work and five levels of curriculum decision-making, the respondents were asked to complete the sentence “The opportunity to become more active in curriculum work is . . .” with one of three endings: “very attractive to me”; “somewhat attractive to me”; and “not attractive to me.” Table 4 records the responses in both absolute numbers and percentages.

TABLE 4  
DISTRIBUTION OF RESPONDENTS BY PERCEIVED ATTRACTIVENESS  
OF CURRICULUM WORK

Perception of Curriculum Work	Absolute Frequency	Percentage of Total
Very attractive	309	24.4
Somewhat attractive	664	52.4
Not attractive	272	21.5
No response	23	1.8
Total	1268	100.0

These data indicate the general degree of interest underlying the preferences displayed in Table 1.

*Discussion*

The purpose of this article was to explore the probable extent of participation of school personnel in curriculum decision-making. The



findings presented in the previous section suggest that participation will be limited in three respects.

First, the participation of school personnel in curriculum decision-making will be limited by the number of individuals who are willing to participate. The general ambivalence of school personnel toward participation is clearly indicated by: (a) primary preference for translating a curriculum into instruction — a kind of curriculum work which is already an essential part of the lives of school personnel and does not expand their role to include broader curriculum concerns, (b) the preference for the school district level of curriculum decision-making — a safe choice which allows greater participation in curriculum decision-making by school personnel who desire it but does not obligate personnel who are disinterested, and (c) the neutral, middle-of-the-road stance of the majority of respondents who found the opportunity to become more active in curriculum work only “somewhat attractive.”

Second, the participation of school personnel in curriculum decision-making will be limited in scope. School personnel are primarily interested in only two kinds of curriculum work. This suggests that they are not prepared at the present time to accept responsibility for the entire range of curriculum decision-making. Whether it is feasible or even possible for school personnel to engage in certain kinds of curriculum work and not in others is a question worthy of consideration.

Third, the participation of school personnel in curriculum decision-making will be limited by characteristics of decision-making at the preferred level of decentralization — the school district. Observation of curriculum decision-making in the context of the school district indicates that a relatively small percentage of personnel employed in the schools of the district are involved in curriculum committees at any one time. In addition, participation usually takes place on an on-again off-again basis as committees are formed and later disbanded. As a result of these two factors, participation in curriculum decision-making will take place *in addition to* the basic work of school personnel rather than as an integral part of it.

Although the findings suggest that the participation of school personnel in curriculum decision-making will be limited by the number of participants, the scope of their work, and the failure to integrate curriculum decision-making into their professional lives, the findings cannot be equated with outright rejection. In fact, there is some evidence that school personnel are taking tentative steps toward greater participation in curriculum decision-making, notably: (a) the second preference for creating a curriculum for a subject which moves school personnel from consideration of *how* to teach to consideration of *what* to teach, (b) interest in the school level of decision-making which typically involves total school staffs, and (c) the “very attractive” response of almost 25% of the respondents to the opportunity to become more active in curriculum work.

It is clear, however, that the decentralization of curriculum decision-making will not simply “happen.” Careful planning will be required to increase the limited participation of school personnel projected by these findings. At least three planning possibilities are suggested by the data.

First, the number of potential participants may be increased, not by compelling ambivalent personnel to participate but by capitalizing on the enthusiasm of the 25% who find curriculum work "very attractive." There is evidence that these personnel are most likely to volunteer for local curriculum projects (Anderson, 1975). If the enthusiasm of these volunteers is channeled into projects which have a high chance of success, the resulting publicity may give ambivalent personnel reason to believe that participation in curriculum work can be "very attractive" indeed.

Second, the scope of participation in curriculum decision-making may be increased by focusing initially on the two kinds of curriculum work which school personnel prefer, not only to capitalize on current interest but also because problems may arise during the projects which will lead naturally to other kinds of curriculum work. It is very possible, for example, that the interest expressed in creating a curriculum for a subject resulted from problems encountered in translating into instruction the very broad guidelines of recent curricula prepared at the provincial level. Another example is the realization of teachers involved in Project Canada West that they needed to win support from other people for their projects (Miller & Dhand, 1973), a kind of curriculum work which school personnel initially find unattractive.

Third, the failure to integrate curriculum decision-making into the professional lives of school personnel suggests a closer look at the possibilities offered by relegating curriculum decision-making to the school level. This procedure, sometimes called school-based management, combines responsibility for curriculum and instructional decision-making, staffing, and inservice training with control of the school budget (Henriquez, 1978; South, 1975). Although there are few formal studies on the effects of school-based management, it appears to offer an opportunity for school staffs to perceive curriculum decision-making as an integral part of their work.

In summary, the evidence presented in this article points to a major obstacle to the decentralization of curriculum decision-making — limited participation by school personnel. Fortunately, the data can be used, not to disparage the merits of decentralization, but to build awareness of a potential problem and to prepare adequately for its solution.

#### References

- Anderson, R. M. *The results of teacher initiative in curriculum development: Some empirical findings*. Paper presented at the annual meeting of the Canadian Association for Curriculum Studies, Edmonton, Alberta, June 1975.
- Beauchamp, G. A. *Curriculum theory* (3rd ed.). Wilmette, Illinois: The Kagg Press, 1975.
- Curriculum decision-making in Alberta: A Janus look*. Edmonton, Alberta: Department of Education, June 1974. (Conference proceedings)
- Henriquez, A. J. *School-based management*. Key West, Florida: Monroe County School District, 1978. Mimeographed.
- Johnson, M. The translation of curriculum into instruction. *Journal of Curriculum Studies*, 1969, 1 (2), 115-131.
- Lortie, D. C. *School teacher*. Chicago: University of Chicago Press, 1975.
- Miller, T. W., & Dhand, H. *The classroom teacher as curriculum developer for Project Canada West*. Saskatoon: Saskatchewan Teachers' Federation, 1973.



- Morrison, T. R., Osborne, K. W., & McDonald, N. G. "Whose Canada?" The assumptions of Canadian studies. *Canadian Journal of Education*, 1977, 2 (1), 73-82.
- Olson, J. K., & Kitto, R. *The role of the teacher in curriculum development*. Paper presented at the annual meeting of the Canadian Association for Curriculum Studies, Fredericton, New Brunswick, June 1977.
- Ponder, A. A., & Bulcock, J. W. Friction point rating: A blueprint for selective decentralization in school systems. *The Canadian Administrator*, 1976, 15 (6), 1-6.
- Resolutions Bulletin*, 10 (12). Edmonton, Alberta: The Alberta Teachers' Association, February 29, 1976.
- School and teacher statistics 1975-76*. Edmonton, Alberta: Alberta Department of Education, Student Evaluation and Data Processing Services Branch, May 1976. Mimeographed.
- South, O. *School-based management*. Tallahassee, Florida: Florida Department of Education, 1975. Mimeographed.
- Sullivan, L. M. Urban school decentralization and curriculum development: Views and implications. In I. E. Staples (Ed.), *Impact of decentralization on curriculum*. Washington, D.C.: Association for Supervision and Curriculum Development, 1975.
- Young, J. H. in collaboration with Young, H. C. *The curriculum decision-making preferences of Alberta school personnel*. Research report duplicated by Printing and Duplicating Services. Edmonton, Alberta: The University of Alberta, 1977.

LINDA SCHLOSSER

and

BOB ALGOZZINE

*University of Florida*

## The Disturbing Child: He or She?

*The ratings by classroom teachers of behaviours which are more prevalent in boys than in girls were compared with regard to their relative effect on teachers' attitudes. The results of three random replications indicated that those behaviours more prevalent in boys were rated as significantly more disturbing to teachers. These findings are discussed with regard to their implications within an ecological frame of reference. (Ms. Schlosser is a doctoral student in the Department of Special Education at the University of Florida; Dr. Algozzine is Assistant Professor in the same department.)*

Differences in the behaviours of boys and girls have been the subject of much research (Brophy & Good, 1974; Maccoby, 1966; Meyer & Thompson, 1956). Interest in sex differences may have been promoted by bits of folk wisdom such as the following:

What are little boys made of?  
Frogs and snails  
And puppy-dogs' tails,  
That's what little boys are made of.  
What are little girls made of?  
What are little girls made of?  
Sugar and spice  
And all that's nice,  
That's what little girls are made of.

(Baring-Gould & Baring-Gould, 1962, p. 176)

While this characterization has an interesting appeal, no research evidence to support its contentions could be found. The prevalence of other behaviour problems in school-aged children has been studied extensively.

In an investigation of 967 kindergarten through third grade children from Minneapolis-St. Paul schools, Rubin and Balow (1971) found that educationally defined behaviour problems (i.e., inability to meet school demands) were present in 41% of the children participating in their study. Their results suggested that teachers accept a relatively narrow range of



behaviours and that deviation beyond that range is seen as cause for outside intervention. In all areas in which results were reported differentially by sex, boys consistently outnumbered girls; that is, they were rated as having repeated more grades, as having more attitude and behaviour problems, and as having received more special services.

In a comparative study of behaviour patterns of learning disabled, emotionally disturbed, and average children, McCarthy and Paraskevopoulos (1969) reported that boys outnumbered girls 9:1 in the learning disabled sample and 8:1 in the emotionally disturbed sample. These results are consistent with those reported by Reinert (1976) who suggested that boys identified as emotionally disturbed by school personnel outnumber their female counterparts six or seven to one and are eight to nine times as likely to be in a special class.

Behaviour problems associated with subpopulations identified as disturbed have been shown to be more prevalent in boys than in girls. Werry and Quay (1971) studied the prevalence of behaviour symptoms commonly associated with emotional disturbance in a sample of 926 boys and 827 girls aged five to eight. Their results indicated that while symptoms of psychopathology were quite prevalent overall in the five- to eight-year-old population, boys tended to exhibit more acting out or disruptive behaviours than girls. The total number of symptoms per child was significantly ( $p < .01$ ) higher for boys ( $\bar{X} = 11.4$ ) than for girls ( $\bar{X} = 7.6$ ). Werry and Quay concluded that there was little doubt that not only did boys have more psychopathological symptoms but the "connotative sense of most of the symptoms commoner in boys represents 'badness'" (p. 142). It was further concluded from this study that boys appear to have a higher rate of disorder and are more "at risk" in our society than girls.

LaPouse and Monk (1964), in a study of 482 school children aged six to twelve who attended schools in Buffalo, New York, reported that generally, extreme scores on behavioural deviations were higher for boys. Statistical significance ( $p < .05$ ) was also reported for boy-girl differences in bedwetting, masturbation, overactivity, temper loss and teacher complaints. Similar results were reported by Schultz, Salvia and Feinn (1974), who assessed prevalence of behavioural symptoms in rural elementary school children in east central Illinois. Using the *Behavior Problem Checklist* (Quay & Peterson, 1975), the authors report that of the 55 problem behaviours on the checklist, 36 symptoms were significantly more prevalent in boys. The overall results appeared to indicate that boys are more conduct disorder prone and demonstrate more inadequate or immature behaviour than girls, who appeared to exhibit more neurotic behaviours. Schultz et al. concluded that their study confirmed the results of Werry and Quay (1971) (i.e., that boys are more "at risk" than girls).

Adams and LaVoie (1974) studied the effect of student's sex, conduct, and facial attractiveness on teacher expectancy. Teachers who participated in the study were asked to complete an evaluation form on a Student Progress Report to which a colour photograph was attached and about which fictitious information was given. Boys were rated lower than girls on attitudes toward school, work habits and personal attitudes. The authors concluded that boys are more difficult to control and are more resistant to

the school regime; they further suggested that “this resistance may be a function of the feminizing influence in the classroom during a time when the boy is actively striving to sex type to the appropriate masculine role” (p. 82).

Sex differences in the distribution of teacher approval and disapproval among sixth-grade children were investigated by Meyer and Thompson (1956). Teacher-pupil interaction within three classrooms was recorded for a total of thirty hours per classroom over an entire school year. Results supported the hypothesis that boys receive a larger number of dominative or disapproval evaluations from their teachers. These differences were interpreted as counter-aggression to the greater expression of aggression by boys.

Harari and McDavid (1973) studied the effects of name stereotypes and teacher expectations, using 80 experienced elementary teachers and 80 female freshmen and sophomores in college as subjects. The effects of name stereotypes were shown to be clearer for boys and this stereotype bias was greater for experienced teachers than for inexperienced college sophomores. Conclusions were drawn concerning the possibility that teachers accumulate these stereotypical expectations and biases over time with their training and experience as teachers.

Investigations have recently demonstrated that the same behaviours which comprise the Behavior Problem Checklist (Quay & Peterson, 1975) represent behaviours which are “disturbing” to teachers in working with children. Algozzine (1977) has derived four dimensions of “disturbingness” from the Quay and Peterson checklist: Factor I, interpreted as a general social immaturity dimension, is perceived as troublesome both for the rater and the child, and is viewed as a precursor of more serious disturbances; Factor II, which is the “social defiance” dimension, is seen as disturbing because it represents serious breaches of the expected social order; Factor III, or “motorically disturbing” traits, is viewed as organically determined and not susceptible to self-control; and Factor IV can be identified as “socialized delinquent” behaviours which are bothersome and often associated with legal transgressions. Results of the investigation indicated that behaviours represented in Factor II (social defiance) were more disturbing than those in Factors I, III and IV. Herr, Algozzine and Eaves (1976) have shown that the disturbing nature of these behaviours can be reduced by intensive exposure to the behaviours during a summer camp experience.

The purpose of the following investigation was to determine if those behaviours which are more prevalent in boys are seen as more disturbing to teachers than are those behaviours which are more prevalent in girls. It was hypothesized that there would be no differences in teachers’ ratings of the relative “disturbingness” of three sets of behaviours.

### *Procedure*

#### *Subjects*

Three groups of teachers were randomly selected from the normative sample utilized in the development of the Disturbing Behavior Checklist (Algozzine, 1977). All of the participants were experienced regular classroom



teachers from a large metropolitan school district. Each group of thirty teachers represented a random replication sample for this investigation. All participants had responded to the 55 items of the Disturbing Behavior Checklist (Algozzine, 1978) indicating how disturbing the behaviours were in working with children. The scaled values for this investigation ranged from 1 (not very disturbing) to 5 (very disturbing) and were used to determine the relative disturbingness of selected behaviours.

### *Method*

An analysis of the results of the major prevalence studies which utilized the Behavior Problem Checklist (Schultz, Salvia & Feinn, 1974; Werry & Quay, 1971) indicated that thirty-two behaviours were more prevalent ( $p < .01$ ) in boys than girls; three behaviours were more prevalent in girls ( $p < .01$ ); and fourteen behaviours were neither more prevalent in boys nor girls. The behaviours more prevalent in boys tended to be the more acting-out and aggressive as well as immature ones (i.e., destructiveness, disobedience). The girl-prevalent behaviours tended to be the more neurotic and withdrawn in character (i.e., behaves like a little adult, shyness, bashfulness). Behaviours prevalent neither in boys nor girls showed a wide range of reactions such as lack of emotional reactivity, self-consciousness, jealousy over attention paid to other children, and truancy from school.

Means of "disturbingness" were calculated because of the large differences in the number of behaviours in each of the three categories. These means were obtained by summing the teachers' responses to the 32 boy-prevalent behaviours and then dividing that total by the number of boy-prevalent items (i.e., thirty-two). The same procedure was followed to obtain means for each teacher's ratings of the girl-prevalent and neutral-prevalent (neither boy- nor girl-prevalent) behaviours. Three means were calculated for each teacher: one for the "disturbingness" of boy-prevalent items, one for the "disturbingness" of girl-prevalent items, and one for the "disturbingness" of the neutral items.

Three separate (i.e., one for each replication) within subjects analyses of variance were performed to assess the relative difference between the obtained means. The level of rejection was set at .01 and further restriction for significance was imposed in that any obtained difference between the means of less than 0.5 units was to be considered trivial. This additional criterion of one-half the unit difference between scale points was seen as necessary and appropriate due to the investigative nature of this study.

### *Results*

The means and standard deviations for each replication sample's responses to the disturbingness of the prevalent behaviours are presented in Table 1. The analysis of variance summary tables are presented in Table 2. In each of the analyses, significant main effects ( $p < .01$ ) were indicated. Since the assumptions regarding homogeneous variance-covariance matrices were not retained, all follow-up analyses to determine significant mean differences were carried out using the individual variances and covariances for each replication sample. In each case, the behaviours more prevalent in boys were viewed as significantly ( $p < .01$ ) more disturbing than those more prevalent in girls; obtained differences between the means

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR DISTURBINGNESS  
OF BEHAVIOURS

Behaviours more prevalent in boys	Behaviours more prevalent in girls	Behaviours neither more prevalent in boys nor girls
<u>Sample One</u>		
$\bar{X} = 2.9$	$\bar{X} = 2.3$	$\bar{X} = 2.6$
$s = 0.6$	$s = 0.8$	$s = 0.6$
<u>Sample Two</u>		
$\bar{X} = 3.0$	$\bar{X} = 2.5$	$\bar{X} = 2.7$
$s = 0.6$	$s = 0.7$	$s = 0.7$
<u>Sample Three</u>		
$\bar{X} = 3.0$	$\bar{X} = 2.4$	$\bar{X} = 2.7$
$s = 0.5$	$s = 0.7$	$s = 0.6$

Note. Score range was: 1-not very disturbing to 5-very disturbing.

TABLE 2  
ANALYSIS OF VARIANCE SUMMARY TABLES FOR THE  
DISTURBINGNESS OF PREVALENT BEHAVIOURS

Source	MS	df	F
<u>Sample One</u>			
Between subjects			
Error	0.996	29	
Within subjects			
Prevalence	2.397	2	11.306*
Error	0.212	58	
<u>Sample Two</u>			
Between subjects			
Error	0.821	29	
Within subjects			
Prevalence	1.432	2	7.232*
Error	0.198	58	
<u>Sample Three</u>			
Between subjects			
Error	0.760	29	
Within subjects			
Prevalence	2.306	2	11.646*
Error	0.198	58	

\* Significant at .01 level.



were greater than the 0.5 unit difference criterion in each case. The neutral items were not seen as more or less disturbing than the girl- or boy-prevalent items.

### *Discussion*

Significant differences were found between the “disturbingness” of behaviours which were more prevalent in boys than those more prevalent in girls. These results suggest that those behaviours which tend to be more characteristic of boys (Schultz, Salvia & Feinn, 1974; Werry & Quay, 1971) are viewed as more bothersome to teachers than are the behaviours characteristic of girls. If boys exhibit more of the behaviours which are bothersome to teachers (Schultz et al., 1974), then their behaviour is more likely to be viewed as more disturbing and of more concern than that of girls.

Ecological theorists have suggested that emotional disturbance may result from an interaction between the behaviours of a child and the effect of the behaviours on individuals in the child’s ecosystem (Rhodes, 1970; Swap, 1974). Within such a model, behaviour is viewed as “disturbing” rather than “disturbed”; equal emphasis is thus placed on both the child and individuals with whom the child interacts. The implications for intervention from an ecological perspective stress the necessity of reducing the frequency of the child’s deviant behaviour *as well as* altering other individuals’ attitudes toward that behaviour (Algozzine, 1977). Such a model suggests that matching of teachers with students may be beneficial and productive. If children who exhibit certain bothersome behaviours are matched with teachers who are tolerant of those behaviours, perhaps “disturbance” in an ecological sense could be avoided.

The strategy of training teachers to be aware of their differential behaviour toward students who exhibit interpersonal differences has shown some promise in changing teacher behaviour effectively (Good & Brophy, 1974). An extension of this concept might include teachers indicating to boys that their behaviour is bothersome and then offering constructive, acceptable less disturbing alternatives to them. This type of mutual “ecological adaptation learning” seems warranted and positively supported by the results of this investigation.

The present study implies that boys exhibit more disturbing behaviours than girls and that teachers should acknowledge and deal with this difference so that it does not become a source of detrimental ecological imbalance within the school environments. Through such a constructive approach to behaviour change and acceptance, perhaps it will be possible to optimize the school experiences of more children.

### *References*

- Adams, G. R., & LaVoie, J. C. The effect of student’s sex, conduct, and facial attractiveness on teacher expectancy. *Education*, 1974, 95, 76-83.
- Algozzine, B. The emotionally disturbed child: Disturbed or disturbing? *Journal of Abnormal Child Psychology*, 1977, 5, 205-211.
- Algozzine, B. Manual for the Disturbing Behavior Checklist. University of Florida, 1978.

- Baring-Gould, W. S., & Baring-Gould, C. *The annotated Mother Goose*. New York: Bramhall House, 1962.
- Brophy, J. E., & Good, T. L. *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart & Winston, Inc., 1974.
- Good, T. L., & Brophy, J. E. Changing teacher and student behavior: An empirical investigation. *Journal of Educational Psychology*, 1974, 66, 390-405.
- Harari, H., & McDavid, J. W. Name stereotypes and teacher expectations, *Journal of Educational Psychology*, 1973, 65, 222-225.
- Herr, D., Algozzine, B., & Eaves, R. Modification of biases held by teacher trainees toward the disturbingness of behavior. *Journal of Educational Research*, 1976, 69, 261-264.
- LaPouse, R., & Monk, M. Behavior deviations in a representative sample of children — variation by sex, age, race, social class and family size. *American Journal of Orthopsychiatry*, 1964, 34, 436-446.
- Maccoby, E. *The development of sex differences*. Stanford, Calif.: Stanford University Press, 1966.
- McCarthy, J. & Paraskevopoulos, J. Behavior patterns of learning disabled, emotionally disturbed and average children. *Exceptional Children*, 1969, 36, 69-74.
- Meyer, W. J., & Thompson, G. G. Sex differences in the distribution of teacher approval and disapproval among sixth-grade children. *Journal of Educational Psychology*, 1956, 47, 385-396.
- Quay, H., & Peterson, D. Manual for the Behavior Problem Checklist, Mimeographed, 1975 (Available from Dr. Quay, University of Miami).
- Reinert, H. *Children in conflict*. St. Louis: Mosby Co., 1976.
- Rhodes, W. C. A community participation analysis of emotional disturbance. *Exceptional Children*, 1970, 36, 309-314.
- Rubin, R., & Balow, B. Learning and behavior disorders: A longitudinal study. *Exceptional Children*, 1971, 38, 293-299.
- Schultz, E., Salvia, J., & Feinn, J. Prevalence of behavioral symptoms in rural elementary school children. *Journal of Abnormal Child Psychology*, 1974, 1, 17-24.
- Swap, S. M. Disturbing classroom behaviors: A developmental and ecological view. *Exceptional Children*, 1974, 41, 163-172.
- Werry, J., & Quay, H. The prevalence of behavior symptoms in younger elementary school children. *American Journal of Orthopsychiatry*, 1971, 41, 136-143.



A. HAMALIAN

Concordia University

## Pupil Control Ideology: Comparative Perspectives — United States and Canada

*The conceptual framework on which this study is based emphasizes the fact that the public school is an organization which does not have control over the selection of clients. Therefore, teachers are dealing directly with non-selected clients and can see themselves as directly responsible for control. The Pupil Control Ideology Scale (PCI) is based on this framework. This scale was used to assess the PCI level of 202 teachers in central and suburban English-language public schools around Montreal in a study conducted during the winter of 1977. The study explored the relationship between PCI and the sex of the teacher, level of teaching (preschool, elementary, secondary), as well as level of school bureaucracy based on the Moeller scale.*

*In line with the results of similar studies from the United States, the results of this study indicated that male teachers were more custodial than female teachers, secondary school teachers were more custodial than elementary school and preschool teachers, and that PCI scores increased with increasing bureaucratization levels. However, on the whole, lower means of PCI were obtained for the group of teachers in this study as compared to the results available from the United States. (Ms. Hamalian is Chairman of the Department of Education at Concordia University.)*

### *The Problem*

Social control is a critical element of group life for establishing and maintaining social order. In society certain institutions assume the responsibility of developing social control in people through a special process of socialization. The family, the school, the church, prisons, and mental hospitals strive to mold the personality of their members to fit the social order acceptable at a particular time in a given society. However, it is true that each of these social institutions may develop control differently and may use different means and mechanisms to achieve this aim.

Public schools as social organizations share certain characteristics with some other institutions of society. Like prisons and mental institutions,

schools do not have much of a choice in the selection of their clientele. On the other hand, the clientele does not have much of a choice either, in the selection of a particular school (Carlson, 1964). Therefore, these organizations and their personnel are confronted with clients who have little or no desire to engage the services offered by these organizations, thus creating the need to control the clientele on the part of the staff. It should be noted, however, that at least two important differences exist in the case of the schools as compared to other organizations. First, prisons and mental institutions are coercive organizations, while schools are normative organizations (Etzioni, 1961). Second, prisons and mental hospitals are "total institutions" while the schools are not (Goffman, 1961). Nevertheless, the similarity of not having a choice over the selection process of the client-organization relationship should have important ramifications in the case of the schools as in the other socialization institutions.

In addition to the fact that schools are organizations which do not have a choice in the selection of their clientele, some of the means and mechanisms involved in the proper functioning of schools are significant in this respect, as well. For example, since the primary task of the teacher is to define and evaluate student levels of achievement, control of most of student behaviour is inevitable and expected in one form or another (Clarke & Hunka, 1977).

Gilbert and Levinson (1957) developed a classification of client control ideology which has been adopted for use in the study of pupil control ideology in public schools in the United States (Willower, Eidell & Hoy, 1967). At one extreme of their conceptual continuum is the "Custodial" orientation which characterizes rigidly traditional schools where the stress is on the maintenance of order, distrust of students and a punitive, moralizing orientation towards pupils. At the other extreme of the conceptual continuum is the "Humanistic" orientation as suggested by Fromm (1948), stressing the importance of the individuality of each pupil, an accepting, trustful view of students, and confidence in their ability to be self-disciplining and responsible.

Furthermore, different schools differ in their policies and orientations concerning control of students. According to Gouldner (1954), the implementation of control through strict rules of reward and punishment may indicate certain tensions within a bureaucratic structure. In this relation, studies which have focused upon schools as social systems have described conflict and problems related to different student subcultures (Coleman, 1961; Willower, Eidell & Hoy, 1967; Gordon, 1957). The interest in these problems has been translated into research and the related literature, by the development of models of custodial and humanistic pupil control ideologies as well as an instrument to measure these different orientations (Willower, Eidell & Hoy, 1967).

Organizational socialization is very important for the secondary socialization process related mainly to occupational and adult roles in an advanced industrial society. Etzioni (1961) states that "Learning of specific skills and role orientations continues with every change of status, in particular with membership in new social units, such as organizations" (p. 142).



Public school teachers go through a double process of organizational socialization: first, at the college level where they learn the ideal norms and patterns associated with the role of a teacher and then through the actual teacher subculture of the particular school where they start their teaching career (Eddy, 1969; Waller, 1967). The following quote from Waller (1967) addresses itself directly to this problem:

The significant people for a school teacher are other teachers, and by comparison with good standing in that fraternity, the good opinion of students is a small thing and of little price. A landmark in one's assimilation to the profession is that moment when he decides that only teachers are important. (p. 389)

The literature on this topic abounds with examples illustrating an orientation toward a desirable permissive control ideology during the college years while the practice of "discipline" in the public schools is of the more authoritarian variety (Ausubel, 1961). Willower, Hoy and Eidell (1967) furnish a few examples of the kind of conflict a new teacher experiences in a school:

Newer teachers reported that a major problem was to convince the older, more experienced teachers, that their younger colleagues were not soft on discipline. The older teachers, dominant in the informal structure of the school, seldom hesitated to communicate their views to the younger, newer teachers whom they often thought of as being lax about maintaining sufficient social distance with regard to pupils. Teachers viewed as weak on control had marginal status among their colleagues and others. Situations of high visibility such as the assembly or school library furnished special testing grounds where teachers made valiant efforts to "look good." Thus, in assemblies, some of the most striking performances emanated from the audience. (p. 229)

Different organizational variables of public schools in their relation to pupil control ideology have been analyzed using the Pupil Control Ideology (PCI) instrument devised by Willower, Eidell and Hoy (1967). Using data from the public school system of the United States, different research studies have supported that teachers are higher on PCI than principals and counsellors (Willower, Eidell & Hoy, 1967); that elementary school teachers are less custodial than secondary teachers; that teachers in schools with a clientele from a higher SES category are less custodial than teachers in schools with a clientele from lower SES categories (Barfield & Burlingame 1974); that teacher PCI varies with the community structure in which the schools operate (Andrews, 1973); that PCI varies with the "open" and "closed" climate of schools (Waldman, 1971; Batista, 1973; Appelberry & Hoy, 1969; Drozda, 1972); that PCI is also related to self-acceptance, self-image and self-actualization levels of teachers (Brenneman, 1973; Noll, Willower & Barnette, 1977; Barfield & Burlingame, 1974).

The variable of sex has not been examined specifically. Whatever results exist in the literature in this respect are either by accident of sampling (Hoy, 1968, 1969) or as an afterthought (Budzik, 1971). In the same way, at the level of the school organizational context the level of bureaucracy of the school has not been directly probed in most of the studies available on this topic. Furthermore, the data pool available is drawn mainly from the public school system in the United States (northeastern and southwestern United States and most particularly from Pennsylvania, New York, New Jersey,

and Oklahoma). Some duplication effort has been conducted in Australia (McArthur, 1973) and in Canada (Fritz, 1973; MacMillan, 1973).

The primary purpose of the project on which this paper is based was to generate some data from Quebec and to explore the relationship between PCI and the sex of the teacher, level of teaching (preschool, elementary, and secondary), as well as level of school bureaucracy.

### *The Instruments*

#### *A. The PCI Form*

The Pupil Control Ideology Form (PCI Form) was developed by Willower, Eidell and Hoy (1967) based on literature, their experience in public schools, field notes from previous studies, and the classification of client control ideology proposed by Gilbert and Levinson (1957). The PCI Form consists of twenty statements measured on a five-point Likert-type scale ranging from "strongly agree" (5-point score) to "strongly disagree" (1-point score). The range of possible scores on the PCI Form is between 20 and 100. A high score signifies a custodial attitude toward pupil control and a low score is indicative of a humanistic attitude toward control of pupils.

According to Willower, Eidell and Hoy, split-half reliability coefficients in two samples were .95 ( $n = 170$ ) and .91 ( $n = 55$ ) with application of the Spearman-Brown formula. Teachers judged to be most custodial by their principals had significantly higher ( $p \leq .01$  using t-testing procedures) PCI Form scores than a like number of teachers judged to be most humanistic. Finally, a cross validation was carried out using the same technique described above (based upon principals' judgements of teacher ideology). Using a one-tailed t-test, it was found that there was a difference between the mean scores of teachers judged to be custodial in ideology and teachers judged to be humanistic in ideology, significant at the  $p \leq .001$  level.

#### *B. The Moeller Scale of Bureaucratic School Systems*

Whatever the basic orientation of an individual teacher toward control, this will converge with the effects of the school organization and social system. The concept of bureaucracy provides a tool for identifying the complexities of the interaction of organizational and human antecedents which have their various effects upon teachers. One of the areas explored by this project was the interrelation of pupil control ideology and the level of school bureaucratization. It was therefore necessary to find a suitable measure of bureaucratization. The measure constructed by Moeller (1964) is based on the characteristics of bureaucracy described by Blau (1956). Since most North American school systems are organized along the bureaucracy model, the distinctions drawn among schools must be within a relatively narrow range on a bureaucratization continuum.

Moeller developed a bureaucracy continuum using an eight-item forced choice instrument on which schools could be scaled from type 0 to 8. A school system to which none of the items applied was scored 0 and the school system to which all eight items applied received a score of 8. The coefficient of reproducibility was found to be .93. The method of scaling the data followed by Moeller is based on the procedures outlined by Riley, Riley and Toby (1954) for construction of an object scale, that is, data from



individuals combined to represent collective responses. Because of the limited number of objects used in scale analysis, the interrater reliability was computed by Moeller and reported to be .47. The reason for the low-interrater correlation is given by the author to be the impossibility of taking out the between rater variance which must go into the error term. Since a subject scale of bureaucracy (for judges) indicated systematic differences between the judges, a cross-tabulation of subject and object scores was made and the correlation as estimated by the coefficient of contingency was virtually zero. Although it is indicated by Moeller that only a study conducted on a far larger sample of ratings could verify the unidimensionality which this measure is presumed to possess, for the purposes of this study we chose this measure of school bureaucratization level because it is very short and very easy to administer.

### *The Sample*

Teachers ( $n = 202$ ) in central and suburban English language public schools around Montreal comprised the random sample for this project, conducted during the winter of 1977. Student trainees in a Sociology of Education course at the university helped in distributing and collecting the questionnaires. Of the initial 250 questionnaires, 231 were returned and of these, 29 were not used because of missing information. Of the remaining 202 questionnaires, 42 (20.8%) were completed by preschool teachers, 65 (32.2%) were completed by elementary school teachers, and 95 (47%) were completed by secondary school teachers. Twelve teachers (5.9%) had only one year of teaching experience, another 57 teachers (28.2%) had two to four years of teaching experience, and the remaining teachers (65.8%) had five or more years of teaching experience.

Sixty-three (31.3%) of the teachers were males and 139 (68.8%) were females. Ninety-three (46%) were from suburban Montreal schools and 109 (54%) were teaching in central Montreal schools. Twenty-seven teachers (13.3%) described their schools as low in bureaucratic organizational level as measured on the Moeller scale, 44 teachers (21.7%) taught in relatively low bureaucratic schools, 60 teachers (30%) taught in relatively high bureaucratic schools and 70 teachers (34.6%) were from highly bureaucratic schools. One teacher (.5%) did not answer this part of the questionnaire. The pupil control ideology of these teachers ranged between 30 and 70 as measured on the PCI scale developed by Willower, Eidell and Hoy (1967). More specifically, one teacher (.5%) had a score between 20 and 29; 25 teachers (12.4%) scored between 30-40; 78 teachers (38.6%) scored between 41-50; 80 teachers (39.6%) scored between 51-60; and 18 teachers (8.9%) scored between 61-70.

The ages of these teachers ranged from 20 years to 61 years of age. Thirty-five teachers (17.3%) were younger than 25, 71 teachers (35.1%) were between 26 and 30 years old; 28 teachers (13.9%) were between 31 and 35 years old; 32 teachers (15.8%) were between 36-40; 25 teachers (12.4%) were between 41-50; and 11 teachers (5.5%) were over 51 years of age. All the teachers in the sample had 13 years or more of formal schooling, 55 (27.2%) had credits towards a B.A. degree but not a completed degree, 28 (13.9%) possessed a B.A., 94 (46.5%) had earned credits beyond a B.A. degree, seven teachers (3.5%) had completed M.A. degrees, twelve (5.9%) had earned

credits beyond the M.A. degree, one had a Ph.D., and five others had different graduate diplomas. For 91 (45%) teachers the undergraduate major was in the field of education, for 63 others (31.2%) it was outside the field of education, and for the remaining 48 teachers (23.8%) it was a combined degree program with education as a minor or a major. One hundred and seventy-seven (87.6%) teachers were Canadian nationals, twenty-five (12.4%) were landed immigrants. The personal income of 51 (25.2%) teachers was less than \$10,000 a year, for 101 (50%) teachers it was between \$10,000 and \$15,000, for 33 teachers (16.3%) this income was between \$16,000 and \$20,000 while 17 teachers (8.4%) had incomes over \$20,000.

Fifty-eight (28.7%) teachers indicated their religious affiliation to be the Catholic Church, 68 (33.7%) were Protestants, 49 (24.3%) belonged to the Jewish faith and 10 (5%) belonged to other religious groups while 17 (8.4%) did not have any particular religious affiliations.

Asked whether they would choose teaching as a career if they had to start all over again, 81 (40.1%) answered yes, 55 teachers (27.2%) said they probably would become teachers again, 44 (21.8%) were undecided and 22 (11%) said they would most probably choose another occupation.

### *The Results*

The purpose of the project on which this paper is based, as stated earlier, was to explore the relationship between PCI and the sex of teacher, the level of teaching, as well as level of school bureaucracy, using a sample of teachers from Quebec to generate data comparative to that already in existence on teachers in the United States.

#### *The PCI of Male and Female Teachers*

In this sample, the results (using a two-tail t-test procedure) indicate a mean score of 51.85 for the 63 males as opposed to a mean score of 48.75 for the 139 females ( $p \leq .01$ ).

This result is in agreement with Budzik's (1971) results where female teachers were found to be more humanistic in PCI than their male counterparts. The data of Hoy (1969) were for all female elementary teachers, but at the secondary school level there were 16 male and 29 female teachers in the group. The mean score on the PCI Form of the female elementary school teachers after first year of teaching was 48.75. At the secondary school level a comparison of the mean PCI score of male and female teachers indicated that the respective means after the first year of teaching were 54.28 for the women and 58.56 for the men.

#### *The PCI and Level of Teaching*

The mean PCI Form score for preschool, elementary, and secondary school teachers were compared and the results indicated respective scores of 47.40, 49.36, and 51.00 (A  $p \leq .05$  was obtained only between preschool and secondary groups).

Once again the results are in agreement with the data existing in the literature as described in the previous paragraph (Hoy's results) and as supported by further data from Willower, Eidell and Hoy studies (1967) where the mean score for elementary teachers was found to be 55.3 as



TABLE 1  
SUMMARY OF THE MEAN PCI FORM SCORE FOR TEACHERS  
IN SAMPLE ( $N = 202$ ) BY SEX, LEVEL OF TEACHING,  
AND LEVEL OF SCHOOL BUREAUCRATIZATION

	Male		Female		Total	
	Number	Mean	Number	Mean	Number	Mean
Sex	63	51.85	139	48.75	202	49.22
		$p < .01$				
Preschool	(only one male)		41	47.44	42	47.44
Elementary	14	54.00	51	48.07	65	49.35
		$p < .01$				
Secondary	48	51.37	47	50.61	95	51.00
		n.s.				
Bureaucracy 1-2 <sup>a</sup>	8	48.37	19	46.31	27	46.92
Bureaucracy 3-4	13	48.84	31	49.41	44	49.25
Bureaucracy 5-6	18	52.44	42	49.50	60	50.38
Bureaucracy 7-8	24	54.20	46	48.60	70	50.52

Note. n.s. = not significant at the  $p < .05$  level.

<sup>a</sup> One teacher did not answer the questions on the level of school bureaucracy. Therefore, this part of the analysis has an  $N$  of 201 teachers.

opposed to 62.3 for secondary teachers ( $p \leq .001$ ). Further evidence in the same direction is provided by Barfield and Burlingame (1974) who show a mean PCI Form score of 49.40 for elementary teachers as compared to 52.85 for senior high teachers and 55.65 for junior high teachers ( $p \leq .01$ ). The preschool teacher group in this study was all female with one exception, and this may be a confounding factor. However, the study proceeded to examine male/female differences at the elementary and secondary level. At the elementary level, the mean score for the women was 48.07 and 54.00 for the men ( $p \leq .01$ ). At the secondary school level, the mean for the women was 50.61 and 51.37 for the men. The male/female difference exists at both levels, although it is less significant at the secondary level.

#### *PCI and Level of School Bureaucracy*

The group mean on the PCI Form for teachers who were teaching in schools with very low bureaucracy scores as measured on the Moeller scale (1-2) is 46.02, for the next level (3-4) a mean of 49.25 was obtained, for the higher level (5-6) the mean was found to be 50.38, and for the highest level (7-8) the mean is 50.52. The PCI increases with increasing level of school

bureaucratization. Significance at  $p \leq .05$  level was obtained only between the lowest (1-2) and the highest (7-8) levels of bureaucratization.

Is there a difference in the PCI of male and female teachers at different levels of school bureaucratization? A comparison of the mean PCI score for men teaching at respective bureaucratization levels of (1-2), (3-4), (5-6) and (7-8) gave the following results: 48.37, 48.84, 52.44 and 54.20, indicating a steady increase in PCI from the lowest to the highest level of school bureaucratization. Significance of  $p \leq .05$  was obtained only in a comparison of groups (3-4) and (7-8).

As for females, the respective mean PCI scores at each level of bureaucratization were 46.31, 49.41, 49.50 and 48.60. Again, there was a trend of increasing PCI scores with increasing levels of bureaucratization (except at 7-8 level). The greatest rise in PCI for male teachers occurred between the three higher levels of bureaucratization while in the case of females, the rise occurs between the first two levels of bureaucratization.

### *Discussion*

The conceptual framework on which this research is based emphasizes the fact that the public school is an organization which does not have control over the selection of clients and because teachers are dealing directly with these non-selected clients, they are or see themselves as directly responsible for their control. Furthermore, since pupils at the elementary school level are seen as posing a lesser threat than pupils at the secondary level, there is the expectation that secondary school teachers will be more custodial in their orientation than elementary school teachers. Previous research has supported this expectation (Willower, Eidell & Hoy, 1967; Barfield & Burlingame, 1976). Our data also support this expectation for the total sample of teachers. The PCI Form mean score for preschool teachers (mostly females) is lower than the PCI Form mean score for elementary school teachers which is itself lower than the PCI Form mean score for the secondary school teachers. However, the difference is only significant between the preschool and secondary school groups. Even here the confounding factor may be the sex of teachers but when broken down along the sex factor, the mean score for the secondary school females is comparable to the mean score of all secondary teachers. Therefore, this expectation is supported.

Although there is no significant difference between the PCI of secondary male and female teachers, there is a significant difference between the two groups at the elementary school level. This may be due to the fact that male teachers have been traditionally associated more with secondary pupils and that they do not feel at ease at the elementary school level and therefore are more inclined to emphasize the custodial aspects of their role.

The second point emphasized in this study was the effect of the level of school bureaucratization on the pupil control ideology of the public school teachers. In recent years, many school reforms have claimed that school bureaucracy is the main factor associated with apathy, custodial orientation towards clients, and rigid ritualistic behaviour. School bureaucracy, it is felt, substitutes rigid conformity to rules in the place of sound professional judgement and orientation. However, in spite of these



possibilities, several studies have described different conclusions. Bonjean and Grimes (1970) as well as Kohn (1971) have concluded that workers in bureaucratic organizations are more enthusiastic, more productive, and exhibit more self-direction than workers in non-bureaucratic organizations where they are faced with confusion because of rule diffuseness. Anderson (1968) applied a similar approach to a junior school setting and found that teachers were least likely to perceive threats to their professional integrity where administration established specific rules for teacher behaviour. Furthermore, a study by Moeller and Charters (1966) indicated that teachers in bureaucratic schools had a significantly higher sense of power in relation to decision making than those in less bureaucratic systems.

As far as results of this study are concerned, there is a steady rise in the PCI Form mean score for the sample with increasing level of bureaucratization (46.93; 49.25; 50.38 and 50.53). However, the inclusion of the preschool teachers and the fact that they are all females confounds this finding. It is known, on the other hand, that the mean score for bureaucratization of the preschools is 5 as opposed to elementary schools where females taught (5.6), elementary schools where males taught (7), secondary schools where females taught (4.9), and secondary schools where males taught (4.4). Therefore, except for the case of elementary male teachers, the distribution of the rest of the sample in schools with different levels of bureaucratization was comparable. Considering males and females separately, there is a steady increase in the mean scores with increasing levels of bureaucratization. It seems that only the extreme levels of bureaucratization make a difference, although there are no significant results to support this fact. Therefore, different levels of bureaucratization do not seem to affect the PCI of the teachers in this sample to any significant extent. Teaching level seems to be more important, and the case of males teaching at the elementary school level is the most problematic.

The most interesting finding of this project is the fact that, irrespective of all other factors involved, the teachers had a lower PCI Form mean score than those in the American samples available in the literature. One important factor may have been the fact of continuing education affecting the teachers in this sample, most of whom were actually involved in some sort of college or university programme to upgrade their academic qualifications. This was a result of the "declassification" procedures which were being implemented by the education authorities in Quebec in an attempt to encourage teachers to take retraining and refresher courses at the university level.

In addition to the fact that the results of this study are very tentative ones because of the size of the sample, the instruments used, lack of control variables, and other research procedures, another important factor to remember is that PCI indicates only an orientation on the part of the teacher and not actual behaviour. Research relating Pupil Control Ideology to Pupil Control Behaviour is in process (Noll et al., 1977). The degree of bureaucratization of the school may be more important when considering actual behaviour patterns because of the "pluralistic ignorance" effect demonstrated by Biddle and his associates (1969):

It was discovered that, for such teacher behavior frameworks as "discipline," "watching for cheaters," "supervision," and the like, respondents attributed to

“People in General” and to “School Officials” much more conservative norms than were actually held by those positions. This distorted view . . . was held by teachers, parents, and pupils alike. Only school officials were able to assess the truly liberal norms held by other school officials and they shared in the general ignorance of public opinion on these issues. (p. 168)

The control function of schools has been a central interest in the field of comparative and international education, especially in its meaning at the societal control level (Writings of the founding fathers and many contemporary researchers, for example Eckstein, 1966). Gradually, this interest is focusing at the classroom level where the general orientations are actually being implemented and transmitted. Interesting studies are being reported but these are isolated case studies. Rust (1973) describes differential control patterns exercised by preschool teachers in Los Angeles, London, and Frankfurt. Direct observation is a most informative method but has many shortcomings such as observer bias, long and expensive training for the preparation of competent observers, language barriers, interference in the classroom activities and its implications, just to cite a few examples. The instrument to measure the Pupil Control Ideology and the related Pupil Control Behaviour are reported to be relatively reliable by several researchers working with different school systems in the United States. It is therefore recommended that similar research be encouraged in Canada as well as in countries having comparable school systems.

#### References

- Anderson, J. G. *Bureaucracy in education*. Baltimore: John Hopkins Press, 1968, 1, 127, 135.
- Andrews, B. J. *Relationships between selected community variables and school atmosphere*. Unpublished doctoral dissertation, Rutgers University, 1973.
- Appelberry, J. B., & Hoy, W. K. The pupil control ideology of professional personnel in “open” and “closed” elementary schools. *Educational Administration Quarterly*, 1969, 5 (Fall).
- Ausubel, D. P. New look at classroom discipline. *Phi Delta Kappan*, 1961, 63, October, 25-30.
- Barfield, V., & Burlingame, M. The pupil control ideology of teachers in selected schools. *The Journal of Experimental Education*, 1974, 42 (Summer).
- Batista, D. M. *The relationships of environmental press and the pupil control ideology of teachers in flexibly-modular and conventionally scheduled high schools*. Unpublished doctoral dissertation, New York University, 1973.
- Biddle, B. J. Roles, goals, and value structures in organizations. In W. W. Cooper et al. (Eds.), *New perspectives in organizational research*. New York: John Wiley and Sons, Inc., 1964, 168.
- Blau, P. M. *Bureaucracy in modern society*. New York: Random House, 1956, 28-33.
- Bonjean, C., & Grimes, M. Bureaucracy and alienation: A dimensional approach. *Social Forces*, 1970, 48 (3), 371.
- Brenneman, O. N. *Teacher self-acceptance, acceptance of others and pupil control ideology*. Doctoral dissertation, Pennsylvania State University, 1973 (underway).
- Budzik, J. M. *The relationship between teachers' ideology of pupil control and their perception of administrative control style*. Unpublished doctoral dissertation, University of Michigan, 1971.
- Carlson, R. O. Environmental constraints and organizational consequences: The public school and its clients. In D. E. Griffiths (Ed.), *Behavioural science and educational administration*. Chicago: University of Chicago Press, 1964, 262-276.



- Coleman, J. S. *The adolescent society*. New York: Free Press, 1961.
- Clarke, S. C. T., & Hunka, S. Comparative views on school discipline. *The Alberta Journal of Educational Research*, 1977, 23 (4), 305-316.
- Drozda, D. G. *The impact of organizational socialization of the pupil control ideology of elementary school teachers as a result of the first year's teaching experience*. Unpublished doctoral dissertation, University of Oregon, 1972.
- Eckstein, M. A. Ultimate deterrents: Punishment and control in English and American schools. *Comparative Education Review*, 1966, October.
- Eddy, E. *Becoming teacher*. New York: Teachers College Press, 1969.
- Etzioni, A. *A comparative analysis of complex organizations*. New York: Free Press, 1961.
- Fritz, J. O. *Control ideology and the student in Alberta high schools*. Proceedings of the Fifth Western Canada Educational Administrators' Conference. C.S.A. Monograph No. 4, Council on School Administration. Edmonton: The Alberta Teachers' Association, 1973.
- Fromm, E. *Man for himself*. New York: Farrar and Rinehart, 1948.
- Gilbert, D. C., & Levinson, D. J. "Custodialism" and "humanism" in mental hospital structure and in staff ideology. In M. Greenblatt, D. J. Levinson & R. H. Williams (Eds.), *The patient and the mental hospital*. Glencoe, Ill.: Free Press, 1957.
- Goffman, E. *Asylums*. Garden City, N.Y.: Doubleday and Company, 1961.
- Gordon, W. C. *The social system of the high school*. Glencoe, Ill.: The Free Press, 1957.
- Gouldner, A. Q. *Patterns of industrial bureaucracy*. New York: The Free Press, 1954.
- Hoy, W. K. The influence of experience on the beginning teacher. *The School Review*, 1968, September.
- Hoy, W. K. Pupil control ideology and organizational socialization: A further examination of the influence of experience on the beginning teacher. *The School Review*, 1969, December.
- Kohn, M. Bureaucratic man: A portrait and an interpretation. *American Sociological Review*, 1971, 36, June.
- Macmillan, M. R. *Pupil control ideology and status obeisance of teachers and principals in elementary schools*. Unpublished doctoral thesis, The University of Alberta, 1973.
- McArthur, J. *Teacher socialization: The influence of experience on the pupil control ideology of beginning teachers*. Australia: Monash University, 1973. (Research underway)
- Moeller, G. H. Bureaucracy and teacher's sense of power. *School Review*, 1964, 72.
- Moeller, G. H., & Charters, W. W. Relation of bureaucratization to the sense of power among teachers. *Administrative Science Quarterly*, 1966, 10 (4).
- Noll, R. L., Willower, D. J., & Barnette, J. J. Teacher self-actualization and pupil control ideology-behaviour consistency. *The Alberta Journal of Educational Research*, 1977, 23 (1).
- Riley, M. W., Riley, J. W., & Toby, J. *Sociological studies in scale analysis*. New Brunswick, N.J.: Rutgers University Press, 1954, Ch. 5.
- Rust, V. D. Teacher control in pre-schools of Los Angeles, London and Frankfurt. *Comparative Education Review*, February, 1973.
- Waldman, B. *Organizational climate and pupil orientation of secondary schools*. Unpublished doctoral dissertation, Rutgers University, 1971.
- Waller, W. *The sociology of teaching*. New York: John Wiley and Sons, 1967.
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology*. University Park, Pa.: Pennsylvania State Studies Monograph, 1967.
- Willower, D. J., Hoy, W. K., & Eidell, T. L. The counselor and the school as a social organization. *The Personnel and Guidance Journal*, 1967, November, 6-11.
- Willower, D. J. & Jones, R. C. Control in an educational organization. In J. D. Rath et al. (Eds.), *Studying teaching*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1967.

JOHN W. KEHOE

and

CHARLES UNGERLEIDER

*University of British Columbia*

## The Effects of Role Exchange Questioning on Empathetic Perceptiveness

*Empathetic perceptiveness may be a necessary condition for changing a disposition of intolerance toward cultural diversity to a disposition of tolerance. The purpose of the study was to test the effects of subjects reading passages from the novel, Under the Ribs of Death by John Marlyn, and then being asked role exchange questions. Subjects were pretested with a Dogmatism scale to determine whether there would be interaction between dogmatism and treatment, and posttested with a measure of empathy. No significant treatment or interaction effects were found. There are at least three possible reasons for lack of treatment effect. First, reading novels and being asked role exchange questions do not enhance empathetic perceptiveness. Second, the scale used to measure empathy was not sensitive enough to discern change. Third, societal factors influencing subjects' attitudes may be so pervasive that subjects are not greatly influenced by what happens to them in school. (Dr. Kehoe is Associate Professor of Social Studies Curriculum and Dr. Ungerleider is Associate Professor of Sociology of Education in the Faculty of Education at the University of British Columbia.)*

Because of the time at which it occurs in a person's life and the amount of time it occupies, public schooling should have a significant impact upon a person's tolerance of cultural diversity. Regrettably, too little is known about the school's impact upon a person's tolerance of cultural diversity and the conditions necessary for changing a disposition of intolerance to one of tolerance.

The development of student empathetic perceptiveness might well be a necessary condition for changing a disposition of intolerance toward cultural diversity to one of significantly less intolerance. If students are able to identify empathetically with ethnic diversity, it is possible that prejudicial attitudes and discriminatory behaviour toward ethnic groups could be changed.



Background

Guira (1967) examined the definition of empathy and found that each definition includes the notion of a shared relationship. The sharing and gaining of understanding through some kind of experience seemed to be an essential aspect of empathy. More recently, Hoffman (1975) defined empathy as the involuntary and, at times, forceful experiencing of another's emotional state. It is elicited either by expressive cues which frequently reflect the other's feelings, or by other kinds of cues which convey the affective impact of external events on the other person.

It has also been argued that empathetic perceptiveness is a necessary component of a moral person. Selman (1975) has stated that empathy is an important component of role-taking ability. Role-taking ability has been shown to be a necessary condition for moving to higher stages of moral development (Kohlberg, 1971). Wilson (1972) has argued that empathy is one of eight components necessary for morality.

Research literature concerning the relationship between dogmatism and intolerance does little to clarify the relationship between dogmatism and empathetic perceptiveness. On the one hand, Rokeach (1960) implies that highly dogmatic personalities would most likely identify with persons in positions of authority who are able to wield power. On the other hand, Fellenbaum and Jackman (1961) have shown that highly dogmatic personalities score high on measures of anxiety. Their work implies that highly dogmatic personalities are able to empathize with the anxiety felt by persons who are subjected to overt discrimination.

Research Design

Purpose

The purpose of the research was to test the effects of role exchange questioning on the empathetic perceptiveness of students. The study was designed to discover whether the impact of the treatment upon the empathetic perceptiveness of students was affected by the level of students' dogmatism or by the teacher to whom the students were exposed. The design of the study is summarized in Figure 1.

Timetable Block		Pretest	Treatment		Posttest
			Experimental	Control	
A	R	O <sub>1</sub>	T <sub>1</sub>	T <sub>2</sub>	O <sub>2</sub>
B	R	O <sub>3</sub>	T <sub>2</sub>	T <sub>1</sub>	O <sub>4</sub>
D	R	O <sub>5</sub>	T <sub>1</sub>	T <sub>2</sub>	O <sub>6</sub>
E	R	O <sub>7</sub>	T <sub>2</sub>	T <sub>1</sub>	O <sub>8</sub>
F	R	O <sub>9</sub>	T <sub>1</sub>	T <sub>2</sub>	O <sub>10</sub>
G	R	O <sub>11</sub>	T <sub>2</sub>	T <sub>1</sub>	O <sub>12</sub>

R = Randomization; T<sub>1</sub> = Teacher #1; T<sub>2</sub> = Teacher #2  
O<sub>1,3,5,7,9,11</sub> = Dogmatism Scale  
O<sub>2,4,6,8,10,12</sub> = Empathy Scale

Figure 1. Study design.

### Sample

One hundred and ninety-nine secondary school students from grades eight, nine, ten, and eleven social studies classes participated in the study. These students were randomly assigned to experimental and control groups within six timetable blocks and were pretested with the Dogmatism Scale (Rokeach, 1960).

### Method

Two teachers alternated teaching experimental and control classes, although they considered both experimental and control classes to be “experimental classes.” Each teacher believed that the purpose of the study was to develop the students’ capacity for understanding another person’s feelings. It was explained that the successful development of this capacity should help in the study of history by enriching students’ understanding of the feelings and motives of the persons about whom they study. The duration of the treatment was two class hours.

Each experimental class read excerpts from Marlyn’s (1957) *Under the Ribs of Death* and then were asked role exchange questions about the feelings they would have if they were in the position of the main character in the novel. *Under the Ribs of Death* is the story of the son of an Hungarian immigrant couple growing up in North Winnipeg in Manitoba. The main character of the book, as a boy and a man, is subjected to humiliation and discrimination because of his ethnic background. His name is mispronounced and he is teased because of it. Teachers, librarians, and nurses are amused when they cannot pronounce his name. He is refused a position with an insurance company because of his ethnic background.

The boy wishes that his family would reject the values of their culture and adopt the values of the English culture. Throughout the novel the boy attempts to transcend his background in an attempt to achieve “success” as he believes it is defined by the English culture. Although the protagonist changes his name from Sandor Hunyadi to Alex Hunter and attempts to manifest attitudes and behaviour associated with the dominant English culture, his ethnic origins are repeatedly discovered and ridiculed and he is subjected to persistent discrimination.

The control class read excerpts from W. O. Mitchell’s (1947) *Who Has Seen the Wind* and were then asked irrelevant role exchange questions. For the most part the questions were related to the main character’s relationship with God.

Upon completion of the treatment all subjects were tested with an eleven-item Empathy Scale. The Hoyt reliability estimate reached .60. A sample item from that scale follows, with the correct response indicated.

A French Canadian born in Quebec goes to British Columbia for the first time and listens to an open line program discussing the French language television outlet to be established in Vancouver. Most of the discussion was strongly opposed to the French station. How do you think he would be likely to react to this?

- (a) He probably thinks they are right at present, since it is causing trouble.
- (b) He may notice it at first, but after a while he probably gets used to it and it doesn’t make much difference to him.



- X—(c) He very likely feels hurt by it and perhaps very angry.  
 ——(d) It is hard to know exactly how he would react to such a situation, though with more information one might be able to tell.

### Analysis and Results

The data were computer analyzed in a three by three factorial design using an analysis of variance technique. An analysis of variance of the posttest scores of the experimental and control groups is reported in Table 1.

TABLE 1  
ANALYSIS OF VARIANCE: EMPATHY SCALE

Source	df	M.S.	F	p
Treatment	1;187	6.02	1.04	<.30
Dogmatism	2;187	18.99	3.30	<.03 *
Teacher	1;187	.45	.07	<.77
Treatment × Dogmatism	2;187	7.12	1.23	<.29
Treatment × Teacher	1;187	.24	.04	<.83
Dogmatism × Teacher	2;187	5.50	.95	<.38
Treatment × Dogmatism × Teacher	2;187	.83	.14	<.86

\* High dogmatism significantly lower in empathy.

### Discussion

The lack of treatment effect was surprising because the teachers reported behaviours and comments on the part of the Marlyn treatment classes which suggested treatment effect. Subjects reported they enjoyed the book and asked for permission to take it home and read it. More of the Marlyn books than of the Mitchell books were taken and not returned. The teachers perceived a shift in class comments from descriptions of how Hunyadi was feeling to descriptions of how they would feel if such things were happening to them. Unfortunately no systematic record was kept of those events so that statistical comparisons could be made.

The study was replicated (Kehoe, 1977) with grade twelve subjects and the empathy scale was expanded to twenty-five items. Again no treatment effects were found. A similar study was conducted by Kehoe, Echols and Barnes (1977) using the novel *Thirty Acres* by Ringuet. It is the story of a French Canadian farmer who loses his land when he is old. A modification of the expanded empathy scale was used and no treatment effects were found. No interaction was found between dogmatism and the treatment. However highly dogmatic individuals tended to exhibit less empathetic perceptiveness than did low dogmatic individuals.

There are at least three possible reasons for lack of treatment effect. First, reading novels and being asked role exchange questions does not enhance empathetic perceptiveness of students. Second, the scale used to measure empathy was not sensitive enough to discern change. Third, societal factors influencing subjects' attitudes may be so pervasive that subjects are not greatly influenced by what happens to them in school.

The notion that societal influences upon students may be so pervasive that students' attitudes are not substantially affected by school curricula is both frightening and intriguing. It suggests that curriculum research of this type might profitably investigate the ways in which students perceive and respond to the curricula presented in school. Through detailed investigation it may be possible to discover the inferences students derive from the formal curricula and the ways in which they integrate or fail to integrate these inferences with those they draw from their experiences outside of the school.

#### References

- Fellenbaum, S., & Jackman, A. Dogmatism and anxiety in relation to problem solving: An extension of Rokeach's results. *Journal of Abnormal and Social Psychology*, 1961, 63, 212-214.
- Guira, A. Z. Toward a systematic study of empathy. *Comprehensive Psychiatry*, 1967, 8 (5), 375-385.
- Hoffman, M. L. Empathy, role-taking, guilt and development of altruistic motives. In T. Lickona (Ed.), *Morality: A handbook of moral development and behavior*. New York: Holt, Rinehart and Winston, 1975.
- Kehoe, J. W. The effect of role exchange questioning on student empathetic perceptiveness. Unpublished replication, University of British Columbia, 1977.
- Kehoe, J. W., Echols, F.H., & Barnes, E. The effects of the novel *Thirty Acres* on student empathy. Unpublished paper, University of British Columbia, 1977.
- Kohlberg, L. Stages of moral development as a basis for moral education. In C. M. Beck, B. S. Crittenden, & E. V. Sullivan (Eds.), *Moral education*. Toronto: University of Toronto Press, 1971.
- Marlyn, J. *Under the ribs of death*. Toronto: McLelland and Stewart, 1957.
- Mitchell, W. O. *Who has seen the wind*. Toronto: Macmillan of Canada, 1947.
- Ringuet. *Thirty acres*. Toronto: McLelland and Stewart, 1940.
- Rokeach, M. *The open and closed mind*. New York: Basic Books, 1960.
- Selman, R. Level of social perspective taking and the development of empathy in children: Speculation from a social-cognitive viewpoint. *Journal of Moral Education*, 1975, October.
- Wilson, J. *Practical methods of moral education*. London: Heinemann Educational books, 1972.



## BOOK REVIEWS

REGISTER OF RESEARCH INTO HIGHER EDUCATION IN WESTERN EUROPE 1974-7. Guildford, England: Society for Research into Higher Education, 1977, 302 pp.

This volume is the second such register published by the Society which has its headquarters at the University of Surrey in the United Kingdom. The register, made possible by financial assistance from the Council of Europe, lists 380 projects which were in progress in Western Europe, excluding the U.K., between 1974 and 1977. All the countries of continental Western Europe are represented but the Netherlands leads the way, with West Germany close behind and France, Sweden, and Finland also very active in this field of research.

The register is a model of organization. Projects are classified in ten different sections, as follows: A, General; B, Economics and Manpower; C, Administration; D, Specific Subjects including curriculum; E, Teaching Methods and Educational Technology; F, Staff; G, Students — General; H, Students — Selection and Performance; I, Students — Choice of Subjects and Careers; J, Further and Adult Education. Of these, the most popular areas with researchers, according to the register, were E and H, with A, D, and J also well supported. On the other hand, section F attracted only nine projects in the period under review.

Every project is numbered and all necessary information is given to enable the interested reader to get in touch with the authors of a particular project. Each entry contains a very brief abstract of the research — just a few lines — to indicate its scope and methodology; the source of financing is noted and also the probable date of completion. If the study has been completed, details of any consequent publications are given. The register concludes with an index of researchers and of sponsoring institutions and organizations.

The register, of course, gives no indication of the conclusions reached in any study or of the relative merits of researches in the same field. It makes no judgements; it is purely a bibliographical handbook and, as such, it is very useful to researchers with a knowledge of the relevant languages. Short summaries of some studies, however, are available in English.

D. R. Pugh  
*Department of Educational Foundations*  
*The University of Alberta*

TEACHING CONCEPTS: AN INSTRUCTIONAL DESIGN GUIDE. By M. David Merrill and Robert D. Tennyson. Englewood Cliffs, N.J.: Educational Technology Publications, 1977, 214 pages, \$14.95.

### *Content*

This textbook is designed to provide an integrated, step-by-step guide to the design and development of instructional sequences and materials using one set of “empirically validated procedures” (p. ix) for teaching a single concept or a set of coordinated concepts. In accord with this intent, the authors do not present a description of alternative instructional strategies, a review of the research literature, or a lengthy discussion of the rationale for selecting particular techniques.

A brief introduction to the instructional design model is presented in the Preface where the reader is also introduced to basic terminology. A detailed explanation of the model is discussed in the subsequent eleven chapters of the book. The discussion includes methods for defining concepts, collecting instance pools of examples and non-examples, estimating instance difficulty, preparing diagnostic tests of classification errors, using attribute isolation (stimulus prompting), designing instructional strategies, teaching coordinated concepts, and formative and summative evaluation. The twelfth and final chapter provides a five-page, flowchart summary of the instructional design model. A brief summary of six related research studies performed by the authors and their colleagues is presented in the appendix. A detailed table of contents lists major and minor topics of discussion. A comprehensive subject index is provided. There are 202 pages of text and illustrations.

### *Instructional Design*

The book is exemplary in its use of instructional design techniques. The authors use concise advance and post organizers as well as brief reviews of previously discussed material to link the chapters together as they evolve a logical and sequential model of instructional design. Important points, terminology, definitions, examples and special notes are variously emphasized through the judicious use of standard and italic type, upper and lower case print, bold and thin face letters, diagrams and photographs, as well as an excellent use of space. These methods are well used and complement rather than detract from the text so that the content rather than the device is salient and the reader is not given the feeling of clever contrivance. The book is very easily and pleasurably read and understood.

The discussion is replete with a variety of illustrative materials drawn from various sources including physics, geography, grammar, politics, advertising and the law. This variety of examples not only illustrates the discussion, provides variety and heightens interest, but also exemplifies the authors’ techniques of teaching concepts, and facilitates generalization. Some examples are repeatedly considered and successively developed over several chapters. The progressive development of a familiar example provides continuity from chapter to chapter.

### *Audience*

The authors state that their “empirically validated” (p. ix) techniques for teaching concepts from “the whole curriculum” (p. 199) to “*most* students” (p. 112, 117, 122 and 134) will produce “concept lessons that are better than you have ever previously taught” (p. xii). Merrill and Tennyson further state



that these methods may be used to assist “teachers, instructional developers, curriculum planners, textbook authors and others who are concerned with effective instructional strategies” (p. xi).

However, the validity of these statements must be questioned in view of the limited number and type of students and concepts studied in the research studies reported in the appendix. Five of the six studies cited were carried out on undergraduate students at the Brigham Young and Florida State Universities, while the sixth study was conducted on junior high school students. These studies examined the teaching of meter in poetry, and the identification of adverbs and chemical crystals. No elementary school children were reportedly studied. Most of the examples of techniques used are presented in a manner most suitable to a textbook instructional format. No techniques are discussed regarding the direct instructional involvement of teachers in common elementary classroom situations. Thus, it appears that the techniques might be most suitable to the design of textbooks for secondary and college level students.

Another difficulty relating to the general usefulness of the text arises from the authors’ apparent attempt to design a book that would be both attractive and useful to novice programmers. In a forenote to the text the authors suggest that to some readers “the procedures outlined in the book may appear to be extremely laborious;” in the final pages of the book the authors report that the comment, “no one can possibly use such detailed procedure” (p. 199), is not unusual. Unfortunately, the authors’ attempt to provide a simplified model of instruction of general appeal appears to have led to the elimination of necessary detail. There is very little discussion regarding the many complex issues relating to the selection and sequencing of examples and non-examples, the use of various instructional formats and a variety of prompting, fading, feedback, reinforcement and correction procedures.

Finally, the reader should note that, although many of the techniques designed to teach concepts have been empirically validated with particular populations of students, no evidence is provided to indicate that the use of the instructional model described in this text will teach readers to use these techniques to “produce concept lessons that are better than you have ever previously taught” (p. xii).

### *Conclusion*

In view of the concepts and procedures discussed in the text, the exemplary manner of their presentation, the ease and pleasure with which the material may be read and understood, I would recommend this text to any undergraduate or graduate student interested in instructional design. However, because of the brief overview of some topics, the elimination of other topics and the absence of practice and feedback exercises, the textbook should be used as introductory material to be supplemented with considerable instruction and supervised application. Although the examples and techniques used appear to be most suitable to the design of textbook materials for secondary and college level instruction, preservice elementary school teachers may reap considerable benefit from reading the text, providing that supplementary instruction is available in the use of direct teacher instructional techniques.

David Baine  
*Department of Educational Psychology*  
*The University of Alberta*

ASSESSING STUDENTS: HOW SHALL WE KNOW THEM? *By Derek Rowntree*. London: Harper & Row Ltd., 1977, 269 pp.

Without benefit of preface or introduction, Rowntree approaches his subject, assessing students, as if driven by the urgency of his subsequent question, "How shall we know them?", a crucial question that has not yet been answered. At the outset, he frankly admits that the topic is not unique: it is his approach that is different. His contention that the literature, extensive but narrow, with its emphasis on only one aspect of student assessment, tests and measurement, is not unfounded. In contrast, he presents "a spectrum of assessment" that ranges from the very informal to the highly formal, surveys a multitude of ways in which we can come to know a student, and emphasizes the 'Why?' as well as the 'How?' of assessment.

Rowntree devotes the first three chapters to upholding the claim that his approach to student assessment is broader than most. The discussion centres around the nature, purposes and possible side-effects of assessment. He recognizes that assessment is a part of evaluation but makes a careful distinction between the two and illustrates, as well, the relationship between educational assessment and "other forms of assessment prevalent in our society." To facilitate an orderly examination of the assessment process, he provides a five-dimensional framework, and then proceeds to "put some flesh on that austere framework" by answering each question posed in relation to each dimension.

Rowntree's questioning strategy is commendable. With emphasis on the heuristics of assessment, he can do no less than ask before telling. There is, however, one question that might be asked of him. Why the inconsistency in two chapter titles which are in the affirmative rather than the interrogative?

Beyond any doubt, the reader is offered plenty to 'chew on.' The questions allow him to think ahead and anticipate how the writer may answer each of the questions (why assess? what to assess? how to assess? how to interpret? and how to respond?) or he may squirm uneasily as the multitude of possibilities unfold. It is highly unlikely that many readers will come away feeling entirely satisfied with the limited ways in which they know their students at present.

However, even the most loyal reader may, at times, grow weary as Rowntree grapples with the problems peculiar to each dimension. Moreover, in spite of his good intention in drawing examples from all levels of education and from other countries, too, Rowntree tends to focus upon assessing students at the university level (where probably the need is the greatest) and in British school settings as, for example, in the discussion of the bureaucratic aspects of assessment (pp. 57-65).

In the final chapter Rowntree makes a valiant effort to draw together the major dimensions of student assessment by answering the key question — "How shall we know them?" — posed in the title of his book. It is obvious that the "seventeen modest proposals" he offers have been written from the point of view of a university teacher, although many of the suggestions, if not all, can be applied to students at any level, primary or university. Besides, earlier in the book, Rowntree does point out that all dimensions need not apply to every situation at a given time.

For whom has this book been written? As stated previously, although discussion throughout the book tends to focus upon assessing students at



the university level (Rowntree is himself a senior lecturer at The Open University), the five-dimensional framework for educational assessment provided by the author can be applied at all levels. Whether a teacher of the young or the not so young, a school administrator, a curriculum planner or an expert on evaluation, the reader may be shaken initially by the straightforward questions raised, yet, in the end, be sufficiently reenforced and energized to act in terms of some of the possible solutions offered. Only those who seek a single answer to the problem of student assessment will be disappointed.

Wilma W. Laing  
*Department of Elementary Education*  
*University of Alberta*

CHILDREN IN ENGLISH CANADIAN SOCIETY: FRAMING THE TWENTIETH CENTURY CONSENSUS. *By Neil Sutherland.* Toronto: University of Toronto Press, 1976, 356 pp., \$19.95 (ISBN 0-8020-5340-8).

In an era when the pinnacle of scholarly publishing in Canadian education appears to be a collection of articles or a volume of documents gathered around a central theme, Neil Sutherland's *Children in English Canadian Society* is a laudable and welcome change. Here we have a seminal work upon which scholars may build. Sutherland has commenced the difficult task of explaining the development of the twentieth century Canadian view of childhood. In the process, he clearly demonstrates the superiority of the monograph as the principal vehicle for disseminating the results of scholarly inquiry.

Unfortunately, the work is made exceptionally complex by the inclusion of a mass of detail indicative of Sutherland's meticulous research and the time he has spent collecting his data. This complexity is unfortunate in that it often overshadows the thread of the author's argument, an argument already obscured by the author himself. Sutherland foregoes the use of a clear introduction and relies on title, chapter headings, and a faith in the deductive reasoning of the reader to elucidate the thesis of the work. Paradoxically, this weakness is also one of the great strengths of the work. This mass of detail will give the book the enduring quality of a major reference to which scholars for many years can be expected to turn.

Sutherland begins by discussing the late nineteenth century attitudes toward children and child-rearing that were characteristic of the Atlantic Provinces and Ontario. Children were viewed as miniature adults valuable as economic resources to most families. An increase in economic prosperity accompanied by a steadily rising standard of living, Sutherland argues, led to a growing determination to invest in ways and means of improving the lot of children. It is the development of this determination that Sutherland seeks to define and explain. He identifies the public health movement, the treatment of juvenile delinquents, and the use of the "New Education" for social reconstruction as the characteristic expressions of the embryonic consensus which was to reach maturity in the twentieth century. An examination of each of these characteristics form the next three sections of the book.

The public health movement was, according to Sutherland, the area which witnessed the greatest reform effort on behalf of children in this

particular period of Canada's history. The movement was primarily concerned with a three-pronged attempt at reform in the fields of health instruction and inspection in schools, infant mortality, and feeble-mindedness. Sutherland is quick to point out that these activities in the field of public health were by no means unique to Canada and he places them within a wider international context.

From public health Sutherland turns to an examination of social health and the concern Canadians expressed about the newly identified problem of juvenile delinquency. He neatly demonstrates the conflict faced by reformers who, on the one hand were delighted with their increasing ability to identify potential criminals and, on the other hand were plagued with doubt arising from the fear that their newly won ability was nothing more than the result of an ever narrowing definition of delinquency. There followed a rather curious growth of institutions to cope with this newly defined problem. Such institutions, as Sutherland clearly shows, had a peculiar facility at adaptation, enabling them to flourish despite the ever changing rhetoric associated with them, the Boys' Reformatory at Penetanguishene being a case in point. Finally, he traces the development of public opinion which came to favour improvements in the family itself as being the most effective method of curbing the growth of a potential criminal element in the society.

Sutherland next investigates attempts at reform in the education system. He isolates two distinct yet complementary reform movements: a campaign to make education more practical, and a drive to make education more self-centred. Considerable attention is given to the impact of people such as Adelaide Hoodless and James L. Hughes and the stimulus provided by the Macdonald-Robinson Movement in the sphere of education. Sutherland concludes this section by tracing the influence of these developments upon practical measures taken through educational legislation passed during the second decade of the twentieth century.

In establishing that a consensus is indeed formed in the twentieth century, Sutherland relies on the formation of the Canada Council of Social Work and a speech delivered by Ontario's Provincial Secretary. The former is presented as a practical and unified expression of the three themes developed in the preceding three sections of the book. The latter is intended to demonstrate the extent to which the people of Canada, in general, subscribed to the views expounded by the reformers. Consequently, it would seem that Sutherland argues that politicians provide a reliable measure of public opinion. It is this general consensus, this middle class desire for a better society, that is apparently the major force behind the changing attitudes that Sutherland has attempted to explain.

After reading the book, one is left wondering whether Sutherland has, indeed, clearly demonstrated the development of a Canadian consensus. Throughout the work he constantly refers to the international context of which the events in Canada formed an integrated part. And, in so doing, he raises questions which must surely be considered fundamental to any research directed toward clarifying specifically Canadian developments. For example, he frequently asserts that Canadians constantly assessed the effectiveness of their programmes by comparing them with similar activities in the United States. But why, we are not told. Rather, one begins to wonder if Sutherland's constant use of the adjective Canadian simply in terms of geographical location has led him to believe that these



developments were also intellectually, and not merely geographically, Canadian.

A similarly often used and beguiling term is English Canadian Middle Class. Sutherland defines this term so broadly that one is left wondering if there was any other class at all in existence in Canadian society. This particular difficulty draws attention to an area much in need of study. It is time that Canadian historians gave some attention to the clarification of such vague and rather woolly terms particularly when they are referred to with such conviction and regularity.

Despite all this, Sutherland's work is destined to be a landmark in Canadian history, both as a first in its particular field and as a standard reference text. Its detail, footnotes, and bibliography will be used by students and scholars alike for years to come. What Sutherland has done is to provide a starting point for a dialogue on the history of childhood in Canada.

J. Stewart Hardy  
*Sessional Lecturer*  
*The University of Alberta*

# Contemporary Research from the Ontario Ministry of Education

## **The Effects of Ontario Teachers' Strikes on Students**

An overview for the general reader of the findings reported in *Three Studies of the Effects of Teachers' Strikes*, to be published later. This summary report provides a theoretical framework in which the overall pattern of school learning can be considered and the probable effects of a disruption in schooling identified.   \$2.50

## **School and Community: Principals and Community Schools in Ontario**

An analysis of the school as a community institution — public or separate, urban or rural — and the many facets of the school-community relationship examined.   \$4.00

## **Sex Roles: Biological and Cultural Interpretations as Found in Social Science Research and Ontario Educational Media**

In what areas are sex differences supported by scientific research? In what way do educators reinforce traditional sex role stereotypes? These are just two of the questions explored in this report.   \$3.50

## **The Development of Improved Bases for Forecasting School Age Population Throughout Ontario: A Study of Demographic Components**

In addition to documentation of trends in population, the implications for school age groups are considered. The projection is based on assumptions derived from trends in fertility, mortality, and immigration.   \$4.50

## **A Cost Analysis Model for Programs in French as a Second Language**

The model does not have a single cost, but a range of possible costs related to decisions about the various properties or parameters of the program. A program is defined, and there is full discussion of the factors influencing the cost structure.   \$3.00

**Orders of \$20 and under must be prepaid.**

Order from: **Publication Sales**  
**The Ontario Institute for Studies in Education**  
**252 Bloor Street West**  
**Toronto, Ontario   M5S 1V6**









## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.





# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXV

NUMBER 2

JUNE 1979

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON

# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

Published quarterly in March, June, September, December by the  
Faculty of Education, The University of Alberta

## CONSULTING EDITORS

J. Britton  
*University of London*

J. Calam  
*The University of British Columbia*

M. Connelly  
*The Ontario Institute for  
Studies in Education*

K. De Clerck  
*State University of Ghent*

R. N. Evans  
*University of Illinois at  
Urbana-Champaign*

R. H. Farquhar  
*University of Saskatchewan*

E. Gagné  
*University of Ottawa*

G. Harman  
*University of Melbourne, Australia*

S. Hunka  
*The University of Alberta*

J. W. G. Ivany  
*Simon Fraser University*

D. A. MacIver  
*University of New Brunswick*

L. D. Nelson  
*The University of Alberta*

W. C. Nesbit  
*Memorial University of Newfoundland*

E. Pedersen  
*McGill University*

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

## FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith  
T. P. Atkinson  
N. C. Bhattacharya  
W. T. Fagan

H. W. Hodysh  
W. W. Laing  
R. G. Martin

E. Miklos (Chairman)  
J. W. Osborne  
C. H. Preitz

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436





# The Alberta Journal of Educational Research

Volume XXV, Number 2

June, 1979

## CONTENTS

Achievement and the Prediction of Achievement in English First and Second Language Children .....	61
<i>B. Panunto and D. White</i>	
Facilitating Reading Comprehension Through Text Structure Manipulation .....	68
<i>J. M. Mason and J. R. Kendall</i>	
School-Centered Community Conflict: The Holdeman Mennonite Case in Alberta .....	77
<i>M. Assheton-Smith and K. Toohey</i>	
A Study of Moral Education in Surrey, B.C. Secondary Schools .....	89
<i>D. M. Williams</i>	
The Effects of Attentional Focus on Graphic Discrimination .....	103
<i>E. Shaw</i>	
Using Qualitative Data in Formative Evaluation .....	117
<i>G. S. Aikenhead</i>	

## BOOK REVIEWS

<i>Research Into Teaching Methods in Higher Education,</i> Ruth M. Beard, Donald A. Bligh and Alan G. Harding .....	130
Reviewed by A. MacKay	
<i>Gladly Would He Teach,</i> John W. Chalmers .....	131
Reviewed by W. J. Bennett	
<i>The Process of Thinking,</i> Marc Belth .....	132
Reviewed by F. N. Walker	

LIST OF BOOKS RECEIVED .....	135
------------------------------	-----

FACULTY OF EDUCATION  
*The University of Alberta*





BRENDA PANUNTO

and

DONNA WHITE

Concordia University

## Achievement and the Prediction of Achievement in English First and Second Language Children

*The purpose of this study was to examine the achievement of Italian first language children in grade one and the predictive relationship between measures of ability and achievement. Subjects were 29 monolingual English-speaking children and 59 English second language children attending grade one classes in middle class areas. The Peabody Picture Vocabulary Test and the Raven Matrices had been administered to all subjects while in kindergarten. In May of the first-grade year, the subjects were given the Metropolitan Achievement Test. No significant differences were found between English first and second language groups on the achievement subtests: the second language children obtained average scores. For Italian children, the best predictor of achievement was the Peabody test given in November of the kindergarten year. (Mrs. Panunto is a Consultant to the Montreal Catholic School Commission; Dr. White is Associate Professor of Education and Psychology at Concordia University.)*

Very little research is available on the school achievement of children for whom English is a second language. The studies which do exist (Ashby, Morrison & Butcher, 1970; Bhatnagar, 1970; Cohen, 1970; Little, Mabey & Whitaker, 1967; McFie & Thompson, 1970; Toronto Board of Education, 1965) indicate that such children achieve at below average levels, tend to drop out of school early, and have social and emotional adjustment problems. These results cannot be accepted without qualification for several reasons. First, these studies utilized children of many different cultural and language backgrounds. Cultural assimilation and acceptance in the school system also varied. English second language status was often confounded with low socioeconomic level. However, even under optimal conditions, it would seem reasonable to hypothesize that children for whom English is a second language would not achieve at average levels as early as grade one simply because they lack English verbal skills. One purpose of the present study was to test this hypothesis utilizing a group of Italian first language subjects

whose English vocabulary tested at the beginning and end of kindergarten was below average (White & Panunto, 1978).

The second purpose of this study was to investigate ability measures as predictors of achievement in English second language children. Achievement in school is dependent on a multitude of factors including verbal and nonverbal abilities. Research has shown that for monolingual English-speaking children, positive correlations between measures of ability and later school achievement hold over a wide range of age, intellectual, and socioeconomic levels (Dokecki, Frede & Goutney, 1969; Dudek, Goldbert, Lester & Harris, 1969; Henderson, Butler & Goffeney, 1969; Hirshoren, 1969; Lesser, Schoeninger & Bridges, 1970; Meyers, Attwell & Orpet, 1968; Mussen, Dean & Rosenbert, 1952; Panther, 1967; Plant & Southern, 1968). Both the Peabody Picture Vocabulary Test (Dunn, 1965) and the Raven Coloured Progressive Matrices (Raven, 1962, 1965) utilized in the present study have been found to correlate with later school achievement (Lesser et al, 1970; Meyers et al, 1968; Panther, 1967; Plant & Southern, 1968). There is some conflict in the literature over whether verbal or nonverbal skills are more effective predictors of achievement in monolingual subjects. Similar significant correlations have been found between both verbal and performance measures and achievement scores. There is some indication that composite measures yield the most powerful predictions for this group (Dokecki et al., 1969; Dudek et al., 1969).

It is difficult to speculate about the relative effectiveness of English verbal skills and nonverbal skills as predictors of achievement in English second language children. There is some indication that both verbal and nonverbal measures predict achievement in English second language children (Cooper, 1958; Phillipus, 1968; Rattan & MacArthur, 1968). However, no studies utilizing kindergarten-grade one subjects could be found.

### *Method*

#### *Subjects*

The subjects were 97 children drawn from 11 schools in the English sector of the Montreal Catholic School Commission in the fall of 1975. In order to control for socioeconomic level, the schools were selected from middle class areas, and inner city schools were eliminated. From the kindergartens of the 11 schools, all available (33) monolingual English-speaking children were included in the control group. A random sample of 64 English second language children was selected for the experimental group. The experimental subjects were all of Italian origin. Children included in this group were those whose parents spoke Italian and little or no English in formal and informal contacts with the schools, and children who, according to their teachers, spoke Italian spontaneously among themselves. Given these criteria, it was inferred that Italian was the language spoken at home and that the experimental children spoke and understood Italian.

All teachers and principals described the Italian families as culturally assimilated in the sense that the parents wanted their children to learn English and to do well in school. In addition, the population of the selected schools was 75-90% Italian, thus the discomfort from cultural differences was minimized.

The mean age of both groups of children in May of 1976 was 6.1 years.



Procedure

The Peabody Picture Vocabulary Test (Dunn, 1965) was administered in November and May of each child's kindergarten year. Alternative forms of the test were used and the order of administration was counterbalanced. The English second language children performed at below average levels in both November and May, while control subjects performed at an average level. The experimental groups scored significantly lower than the control groups on both Peabody tests (White & Panunto, 1978).

The Raven Coloured Progressive Matrices (Raven, 1962, 1965) was administered individually in May of the kindergarten year. Both experimental and control subjects performed at similar, average levels on the Raven (White & Panunto, 1978).

In May of the grade one year, the Metropolitan Achievement Test, Primary 1 Battery — Form F (Durost, Bixler, Wrightstone, Prescott & Balow, 1971) was administered in groups to 88 of the original 97 subjects still in the school system.

Results

Achievement Tests

The mean scores obtained by the control and experimental groups on the subtests of the Metropolitan are shown in Table 1. It can be seen that control subjects earned higher scores on all subtests although, with the exception of mathematics, the differences seem to be slight. It should be noted that all of these mean scores represent average performance on the Metropolitan Achievement Test.

TABLE 1  
ACHIEVEMENT MEANS AND *T* TESTS FOR ENGLISH FIRST  
AND SECOND LANGUAGE GROUPS

Achievement Subtests	Groups		<i>t</i> test Between Groups df = 86
	English First Language	English Second Language	
Word Knowledge	23.4	20.8	1.54
Word Analysis	28.9	27.4	.94
Reading	21.3	19.4	1.02
Total Reading	45.1	40.2	1.46
Total Mathematics	50.8	37.6	1.37

In order to determine whether the differences between the mean scores obtained by the control and experimental groups were significant, a series of *t* tests were performed. As indicated in Table 1, there were no significant differences between the groups on any of the subtests. Contrary to our prediction, the Italian children in this study performed at average levels on achievement tests in grade one and were not significantly different from monolingual English controls.

Prediction of Achievement

*Correlations.* It was hypothesized that both verbal and nonverbal scores would correlate with achievement in both English first and second language subjects. In

TABLE 2  
CORRELATION OF ABILITY AND ACHIEVEMENT MEASURES IN  
ENGLISH FIRST AND SECOND LANGUAGE GROUPS

Achievement Subtests	Ability Tests		
	Peabody- Nov.	Peabody- May	Raven
<u>English First Language Group</u>			
Word Knowledge	.071	.315	.386*
Word Analysis	-.132	.063	.286
Reading	.253	.344	.483*
Total Reading	.165	.344	.429*
Total Mathematics	.278	.184	.307
<u>English Second Language Group</u>			
Word Knowledge	.568**	.400**	.264*
Word Analysis	.403**	.338**	.224
Reading	.401**	.302*	.054
Total Reading	.521**	.377**	.169
Total Mathematics	.542**	.471**	.377**

\*  $p < .05$

\*\*  $p < .01$

order to test this hypothesis, a Pearson  $r$  was calculated between each of the three ability scores and each of the five subtest scores of the Metropolitan. As can be seen from Table 2, the Peabody scores are significant predictors of school achievement for the Italian children tested, but are not significantly correlated with achievement measures for control subjects. The Raven scores are correlated significantly with some of the achievement subtests for both groups.

*Step-wise regression.* Since the independent variables have been shown to be intercorrelated (White & Panunto, 1978), a step-wise regression analysis (Nie, Bent, & Hull, 1970) for the English second language group was performed to determine whether prediction could be improved by using several independent variables. The difference between the Peabody Vocabulary scores in November and May was included as a fourth independent variable in this analysis. As can be seen in Table 3, no significant gain in predictive validity was made for English second language subjects by adding data of test scores other than the Peabody Picture Vocabulary Test given in November.

### Discussion

The finding that Italian first language children performed at average levels on achievement tests in grade one forces rejection of the idea that immigrant children will always do poorly in school. When factors such as socioeconomic level, nonverbal ability, and comfort within the school system are controlled, English second language children seem to rapidly overcome verbal deficiencies. It should be noted that since the Italian children represented the majority group in each class, the teachers geared the academic programmes to these children. The teachers in this study reported that there was an emphasis on learning English language skills



TABLE 3  
STEP-WISE REGRESSION ANALYSIS FOR THE  
ENGLISH SECOND LANGUAGE GROUP

Achievement Subtests	Predictor Variables	Multiple R	R <sup>2</sup> Change
Word Knowledge	Peabody-Nov.	.568	.322
	Raven	.575	.008
	Peabody-May	.575	.001
	Difference (Nov.-May)*		
Word Analysis	Peabody-Nov.	.403	.163
	Raven	.416	.010
	Peabody-May	.427	.009
	Difference (Nov.-May)*		
Reading	Peabody-Nov.	.401	.161
	Raven	.408	.006
	Peabody-May	.410	.001
	Difference (Nov.-May)*		
Total Reading	Peabody-Nov.	.521	.271
	Difference (Nov.-May)	.522	.001
	Peabody-May*		
	Raven*		
Total Mathematics	Peabody-Nov.	.542	.294
	Raven	.584	.047
	Difference (Nov.-May)	.605	.025
	Peabody-May*		

\* *F* level insufficient for further computation.

which seems to have resulted in balancing the English and Italian known by the experimental group. It would be of interest to study the programmes of these teachers since no formal second language curriculum plan existed in these schools.

With regard to the prediction of achievement in English second language children, this study provides clear support for the correlation of English verbal ability and achievement and some support for the relationship of nonverbal abilities and achievement. These findings are consistent with the work of Cooper (1958), Phillipus (1968), and Rattan and MacArthur (1968), which utilized older subjects of varying ethnic backgrounds.

The most dramatic finding seems to be the strong correlation of the Peabody Vocabulary Test administered in November and achievement in the English second language group. It is possible to speculate that English verbal skills acquired prior to kindergarten may be instrumental in academic achievement. A study of the factors which lead to the early acquisition of English by some Italian children seems warranted. Variables such as parental support, older siblings who are learning English, and nursery school exposure to English should be investigated.

The findings of the present study regarding prediction of achievement for monolingual control subjects are somewhat inconsistent with the literature, particularly for verbal skills. The lack of significant correlations between Peabody scores and achievement may be attributed to the small size and somewhat unique characteristics of the control group utilized in this study. These characteristics

include the fact that the monolingual children represented a minority group in the selected schools and had academic programmes which were not directed toward optimal development of their reading and mathematics skills.

This paper is based on a master's thesis completed in the Department of Education, Concordia University. The research was partially supported by a grant from Formation de Chercheurs et d'Action Concertée (Minister of Education, Province of Quebec).

### References

- Ashby, B., Morrison, A., & Butcher, H. J. The abilities and attainments of immigrant children. *Research in Education*, 1970, 4, 73-80.
- Bhatnagar, J. *Immigrants at school*. London: Cornmarket Press, 1970.
- Cohen, D. K. Immigrants and the schools. *Review of Educational Research*, 1970, 40, 13-27.
- Cooper, J. G. Predicting school achievement for bilingual pupils. *Journal of Educational Psychology*, 1958, 49, 31-36.
- Dokecki, P. R., Frede, M. C., & Goutney, D. B. Criterion, construct, and predictive validities of the Wechsler Preschool and Primary Scale of Intelligence. *Proceedings of the 77th Annual Convention of the American Psychological Association*, 1969, 4, 505-506.
- Dudek, S. Z., Goldbert, J. S., Lester, E. P., & Harris, B. R. The validity of cognitive, perceptual-motor and personality variables for prediction of achievement in grade I and grade II. *Journal of Clinical Psychology*, 1969, 25, 165-170.
- Dunn, L. M. *Expanded Manual for the Peabody Picture Vocabulary Test*. Minneapolis: American Guidance Services, 1965.
- Durost, N., Bixler, H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. *Metropolitan Achievement Tests, Primary I, Forms F, G, and H*. New York: Harcourt, Brace, and Jovanovich, 1971.
- Henderson, N. B., Butler, B. V., & Goffeney, B. Effectiveness of the WISC and Bender-Gestalt Test in predicting arithmetic and reading achievement for white and nonwhite children. *Journal of Clinical Psychology*, 1969, 25, 268-271.
- Hirshoren, A. A comparison of the predictive validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. *Exceptional Children*, 1969, 35, 517-521.
- Lesser, K., Schoeninger, D. W., & Bridges, J. S. Prediction of first grade performance. *Perceptual and Motor Skills*, 1970, 31, 751-756.
- Little, A., Mabey, C., & Whitaker, G. The education of immigrant pupils in inner London Primary Schools. *Race*, 1967-68, 9, 439-452.
- McFie, J., & Thompson, J. A. Intellectual abilities of immigrant children. *British Journal of Educational Psychology*, 1970, 40, 348-351.
- Meyers, C. E., Attwell, A. A., & Orpet, R. E. Prediction of fifth grade achievement from kindergarten test and rating data. *Educational and Psychological Measurement*, 1968, 28, 457-463.
- Mussen, P., Dean S., & Rosenbert, M. Some further evidence on the validity of the WISC. *Journal of Consulting Psychology*, 1952, 16, 410-411.
- Nie, N., Bent, D., & Hull, C. H. *Statistical Package for the Social Sciences*. New York: McGraw Hill, 1970.
- Panther, E. E. Prediction of first grade reading achievement. *Elementary School Journal*, 1967, 68, 44-48.
- Phillipus, M. J. Test prediction of school success of bilingual Hispanoamerican children, *Eric Report # ED 036577*, 1968.
- Plant, W. T., & Southern, M. L. First grade reading achievement predicted from WPPSI and other scores obtained 18 months earlier. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, 3, 593-594.
- Rattan, M. S., & MacArthur, R. S. Longitudinal prediction of school achievement for Metis and Eskimo pupils. *Alberta Journal of Educational Research*, 1968, 14, 37-41.



- Raven, J. C. *Coloured Progressive Matrices, Sets A, Ab, B*. London: Lewis and Co., 1962.
- Raven, J. C. *Guide to using the Coloured Progressive Matrices, Sets A, Ab, B*. London: Lewis and Co., 1965.
- Toronto Board of Education. *Immigrants and their education*. Toronto Board of Education, 1965.
- White, D., & Panunto, B. Verbal and nonverbal abilities in English first and second language children, *Psychological Reports*, 1978, 42, 191-197.

JANA M. MASON

*University of Illinois at Urbana-Champaign*  
and

JANET ROSS KENDALL

*Simon Fraser University*

## Facilitating Reading Comprehension Through Text Structure Manipulation

*Two experiments were conducted using 9 to 12-year-old readers to determine how they differ from adults in their ability to identify meaningful units in a written passage, how they are affected by sentences of increasing difficulty, and whether they are aided by either shorter sentences or meaningful segmentation of sentences. The first experiment showed that children mark insufficient or inappropriate intrasentence pausal breaks. The second experiment determined that shortened or segmented sentences may increase reading time and, with low ability readers, will improve reading comprehension. There was also an indication that for all children, reading comprehension can be affected by text characteristics even when sentence length is controlled and word length is manipulated systematically. (Dr. Mason is Assistant Professor, Center for the Study of Reading at the University of Illinois at Urbana-Champaign; Dr. Kendall is Assistant Professor in the Faculty of Education at Simon Fraser University.)*

There is evidence that an ability to identify intrasentence units of text is an important aspect of comprehension (Rode, 1974-75; Weinstein & Rabinovitch, 1971) rather than of reading efficiency alone (Coleman & Kim, 1961). Investigators (Oakan, Wiener & Cromer, 1971; Steiner, Wiener & Cromer, 1971) have demonstrated that a distinguishing characteristic of good and poor readers is poor readers' inability to identify and organize these units; thus poor readers' comprehension is significantly inferior to that of good readers. This topic should be pursued, both to learn under what circumstances poor readers are impaired and to determine whether manipulations of syntactic complexity can improve reading comprehension.

A technique for identifying meaningful units in written text was first employed by Johnson (1970) who found a high agreement among skilled readers when they were asked to place slash marks at the points where they would pause when reading a passage aloud. Frase and Schwartz (1977) relied on this procedure to show that separation of sentences into indented, meaningful units reduced the length of time



needed to locate information. Hartley and Burnhill (1976), Cromer (1970), Coleman and Kim (1961), and North and Jenkins (1951) found that text materials that were separated into phrases reduced errors in finding information, improved comprehension scores for poor readers, or led to faster reading.

In research with children, Buswell (1920) found that good comprehenders had longer eye-voice spans than did poor comprehenders. This was confirmed by Levin and Kaplan (1970). Eagan (1973) showed that the number of oral reading pauses decreased from grade 2 to 3 and from poor to good readers even after controlling for decoding skill. These studies indicate that children who are better comprehenders have the longer eye-voice spans, implying an ability to identify and group appropriate intrasentence units.

Based on the eye-voice span research, there is reason to expect age and reading skill differences in an ability to identify intrasentence pausal junctures. The effect would be more important, however, if an improvement in reading comprehension could be obtained from text that was segmented by intrasentence junctures, especially if poor readers in the upper elementary grades should obtain higher reading comprehension scores when sentences are segmented into meaningful units than when left unchanged. That is, finding a greater difference between low and high ability readers on the standard text than on the segmented text, with only low ability readers being helped by the manipulation, would indicate that comprehension can be affected by the *obviousness* of intrasentence units.

In the first experiment reported here, adults and 9- to 11-year-old children were asked to read and segment into pausal junctures a passage which was graded by paragraph from a grade 1 to a grade 9 level of difficulty. This study was intended to determine whether children are able to locate intrasentence junctures. In the second study, fourth graders read a story in a standard or revised format. Children marked the time it took them to read each of four sections of the passage; after reading they answered comprehension questions about the passage. Format effects were expected using children's reading time and comprehension error rate as dependent measures; format effects were also predicted as a function of reading ability.

### *Experiment 1*

The purpose was to determine whether children who could recognize most of the words in a text would be able to segment sentences into meaningful units in a manner similar to that of adults.

### *Method*

Twenty-two college students and 60 children aged 9, 10, and 11 read a nine-paragraph passage. The passage had been normed as an oral reading task; the first paragraph was at a grade 1 level of difficulty and each succeeding paragraph increased in difficulty up to grade 9. Adults and children were asked to place slash marks where they would pause if they were to read the passage aloud; children were also asked to circle words they could not pronounce and stop reading if they circled five or more words in a paragraph. The tasks were carried out in classroom settings. Eight children's responses were omitted from the analysis because they circled several words from or did not complete the 5th or 6th grade level paragraph. Of the remaining children, 8 were 9 years old, 21 were 10, and 23 were 11. All of the children appeared to have made realistic assessments of their ability to read the materials.

Only three of the nine paragraphs were analyzed — those approximating the difficulty levels of grades 4, 5, and 6. The easier passages were not analyzed because adults determined that these sentences contained very few breaks (only 3 in the 21 sentences). The more difficult passages were not analyzed because fewer children completed them and those who did tended to restrict themselves to marking commas. In the three analyzed paragraphs there were 22 intrasentence junctures that were agreed upon by half or more of the college students. These junctures were compared with those produced by the children in terms of hits (slash marks that agreed with adults'), misses (slash marks that disagreed with adults' marks), and proportion of responders who parsed only punctuation marks.

### *Results*

Large differences between adults' and children's judgments were found, particularly in the number of slashes. Children marked only six to nine of the junctures that at least half of the adults had located. Smaller differences appeared in the number of misses primarily because some children relied almost entirely on comma junctures for placement of slashes. Forty-two and 43% of the 10- and 11-year-olds but none of the 9-year-olds and adults marked only commas. With these materials, the 9-year-olds made idiosyncratic judgments, ignoring most commas as well as other appropriate junctures; 10- and 11-year-olds were quite cautious in their judgments; and adults, while agreeing on the essential pausal breaks, tended in addition to mark lesser pause breaks (see Table 1 for summary).

TABLE 1  
TABULATION OF HITS, MISSES, AND AN ACCURACY RATIO  
OF INTRASENTECE PAUSAL BREAKS

Age	Average hits	S.D.	Average misses	S.D.	Hits/ Total response
9	6.5	3.1	8.1	7.9	.44
10	9.0	4.8	3.4	6.5	.73
11	9.2	4.2	3.0	4.9	.75
College	15.3	5.1	6.6	6.3	.70

### *Discussion*

Children cannot readily separate sentences into meaningful units although the ability to do so shows some improvement with age. Nine-year-olds make little use of any consistent rule in marking boundaries; however, it should be noted that two of the passages may have been too difficult for them to read. The greater restraint and accuracy of 10- and 11-year-olds over the younger children and even adults indicates that although the task is understood, there is an insufficient ability to locate pausal boundaries.

It should be noted that because only eight 9-year-olds remained in the sample, differences between them and older children must be interpreted cautiously. The major point, that children make judgments quite different from those of adults, is unaffected by the small number of 9-year-olds.



### Experiment 2

An inability of upper elementary school children to distinguish in sufficient detail intrasentence units might be related to reading comprehension. In the second experiment, this matter is tested by hypothesizing that low ability readers cannot locate intrasentence junctures in text materials but if junctures are made more obvious, comprehension scores ought to be improved and reading rate changed. This is tested by measuring the reading comprehension and reading rate effects of texts materials which are changed specifically to make intrasentence units more apparent.

#### Method

*Subjects.* Ninety-eight fourth grade students from a metropolitan Canadian city who had not been used in the first experiment were tested. Schools and classrooms were selected on the basis of Metropolitan Achievement Test (MAT) score averages which were at or near the 50th percentile.

*Materials.* The three paragraphs that were analyzed in Experiment 1 served as a content base for the construction of systematically varied materials. Four levels of difficulty were created by dividing the paragraphs into 4 sections, each of which contained five sentences; each sentence in the first section contained 18 syllables, the next had 22 per sentence, then 26, and the fourth section had 30. Five multiple-choice comprehension questions were constructed for each of these four levels. Two variations of these materials were put together so that in one, pausal units appeared on separate lines of the page while in the other, four short sentences reflecting these units replaced each original sentence. The complete structure is summarized in Table 2. An example of the materials is shown below.

#### Standard

Dick will be in Grade Five and though he enjoys math he likes art class best.

#### Parsed

Dick will be in Grade Five  
and though he enjoys math  
he likes art class best.

#### Short sentence

Dick will be in Grade Five. He enjoys math. He enjoys art. He likes art class best.

#### Questions from comprehension test

What is Dick's best subject?

(a) math    (b) reading    (c) art    (d) science

*Procedure.* In group testing, children read the passage in one of three format conditions: (1) standard text, (2) each sentence broken into four short sentences, and (3) each sentence divided by meaningful units onto separate indented lines. When they finished each of the four difficulty level sections, they wrote down the time that was displayed on the blackboard (the time was changed every 15 seconds). When they finished the passage, they responded to the vocabulary and comprehension questions which followed. They were instructed to answer the



TABLE 2  
DESCRIPTION OF FORMAT AND INTRAPASSAGE DIFFERENCES

Format condition	Passage section			
	First	Second	Third	Fourth
Short sentence				
Number of lines	9	11	14	15
Number of sentences	20	20	20	20
Number of words per sentence	5.35	7.25	9.30	10.15
Number of syllables per sentence	5.85	5.85	6.40	6.55
Parsed				
Number of lines	11	12	11	12
Number of sentences	5	5	5	5
Number of words per sentence	17	17	17	17
Number of syllables per sentence	18	22	26	30
Standard				
Number of lines	8	8	9	9
Number of sentences	5	5	5	5
Number of words per sentence	17	17	17	17
Number of syllables per sentence	18	22	26	30

questions without rereading the story and were given as much time as they needed to read and to answer the questions.

Results

Before running an analysis of variance, scores from the *MAT* were used to construct 3 approximately equal-sized ability groups (the low group scored between 5 and 20 and the high group was between 30 and 42). Then, *passage format* and *reading ability* were between subjects variables while *passage difficulty* was a within subjects variable. This design was analyzed three times using comprehension error, reading time, and the ratio of reading time over the number of syllables per standard sentence as dependent variables. Mean scores from these analyses appear in Table 3.

*Comprehension error.* With an average error rate of 28%, significant effects were obtained of reading ability,  $F(2,87) = 15.0, p < .001$ , passage difficulty,  $F(3,261) = 23.0, p < .001$ , and a borderline interaction between reading ability and passage format,  $F(4,87) = 2.4, p < .06$ . However, an a priori test was warranted in analyzing the interaction because low ability readers had been predicted to be helped by the experimental formats. *T*-test comparisons of the format conditions indicated significant differences for low ability readers between the standard and short sentence formats,  $t(87) = 4.29, p < .01$ , and between the standard and parsed sentence formats,  $t(87) = 2.0, p < .05$ . No significant *t*-test differences were found for middle or high ability students. The interaction is shown in Figure 1, on the left.

The main effect of passage difficulty indicated, contrary to expectation, that paragraphs containing the shortest words (first section) and longest words (last section) generated the lowest error rates. This will be discussed later.

*Reading time.* The reading time analysis showed significant effects of reading

TABLE 3  
ERROR RATES AND READING TIMES UNDER PASSAGE FORMAT AND  
PASSAGE DIFFICULTY CONDITIONS, AND AS A FUNCTION OF  
READING ABILITY

	Mean		Ratio of Reading time to Number of syllables
	Error rate	Reading time in seconds	
Format			
Standard format	1.51	39.7	.34
Parsed format	1.46	45.3	.38
Short sentence format	1.29	55.5	.34
Passage difficulty (section)			
First	1.00	34.2	.35
Second	1.72	47.1	.39
Third	1.94	50.3	.34
Fourth	1.02	55.6	.33
Reading ability			
High	1.04	38.3	.29
Middle	1.27	42.6	.32
Low	1.95	59.5	.44

ability,  $F(2,87) = 7.1, p < .001$ , passage difficulty,  $F(2,87) = 33.9, p < .001$ , passage format  $F(2,87) = 3.6, p < .03$ , and a passage format by passage difficulty interaction,  $F(6,261) = 3.0, p < .01$ . The easiest passage section took an average of 32 seconds to read, the hardest took 56 seconds. Reading ability (MAT scores) was correlated significantly with reading time under the two experimental conditions at every level of difficulty, but not with the standard format (see Table 4).

TABLE 4  
CORRELATIONS BETWEEN MAT COMPREHENSION AND READING TIME  
AS A FUNCTION OF TEXT FORMAT AND TEXT DIFFICULTY

Format	Text Difficulty				Whole Text
	1 (Easiest)	2	3	4 (Hardest)	
Standard	+.01	+.07	+.04	-.20	-.05
Parsed	-.61*	-.59*	-.53*	-.60*	-.64*
Short sentence	-.58*	-.54*	-.55*	-.59*	-.60*

\*  $p < .01$

The reading ability effect indicated that high ability readers were the fastest readers while low ability readers were the slowest. These differences were maintained over all levels of text difficulty but not uniformly over format conditions. Although the latter interaction was not significant, since a difference had been predicted and correlations indicated a discrepancy, Tukey's test was made to evaluate the effects of format on reading time of high and low ability

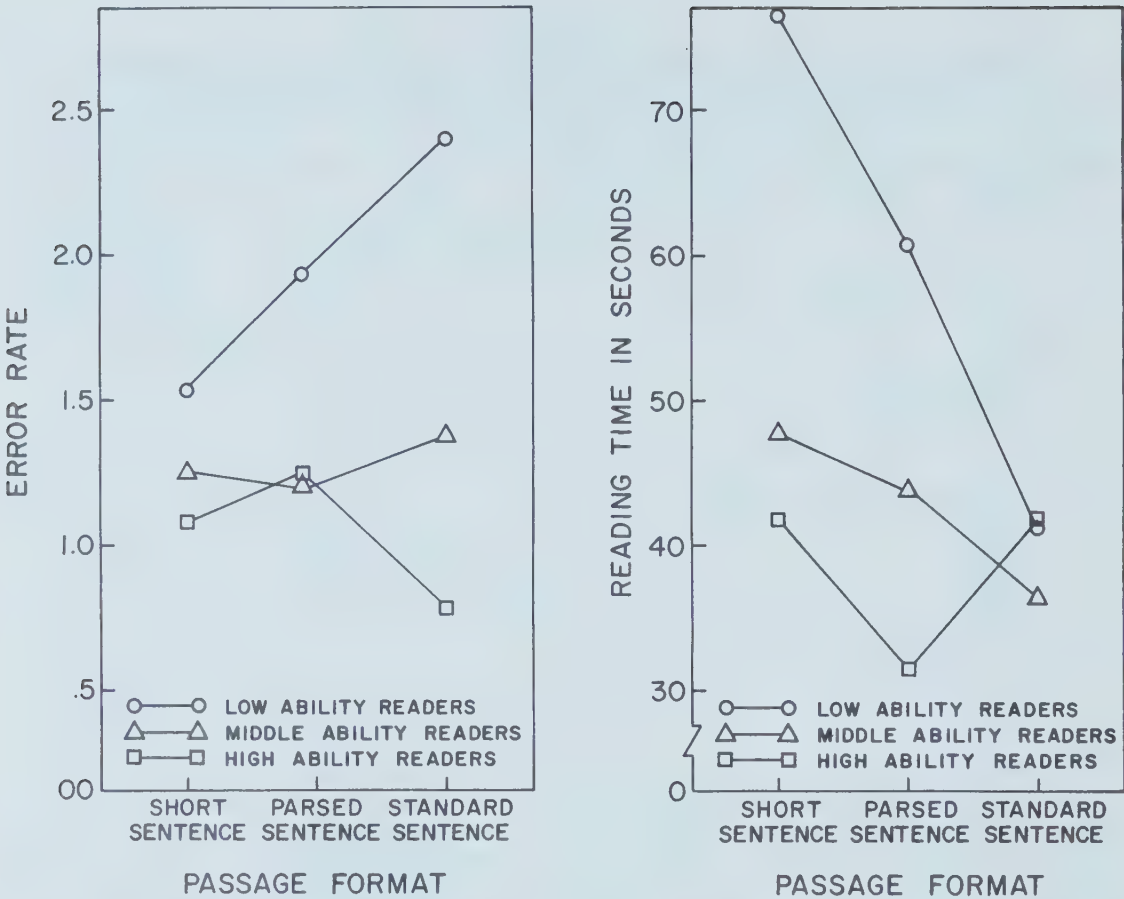


Figure 1. Interactions between reading ability and passage format for error rate and for reading time.

readers. There was a significant difference under the short sentence condition,  $q(3,32) = 4.2$ ,  $p < .05$ , a nearly significant difference in the parsed condition,  $q(3,32) = 3.1$ ,  $p < .10$ , and no difference in the standard condition. These effects are graphed on the right side of Figure 1.

The format main effect determined that the experimental formats increased reading time with the short sentence taking the longest time. The interaction between passage format and passage difficulty showed that as the number of syllables per sentence was increased, more time was spent reading. This was particularly true for the two experimental formats.

In a third analysis using a ratio of time over syllable number, there were neither significant format differences nor a significant interaction with passage difficulty, permitting the interpretation that format differences where time alone was the dependent measure were caused by the short sentence format which contained fewer words per sentence. The passage difficulty effect was still significant in the ratio analysis,  $F(3,261) = 5.6$ ,  $p < .001$ , as was reading ability,  $F(2,87) = 6.7$ ,  $p < .002$ .

Frase and Schwartz (1977) found that appropriately parsed and indented sentences reduced reading time. In this study, parsing did not reduce reading time. This is probably due to the task differences. Frase and Schwartz had people read until information was found rather than read to understand and remember information. Also, the amount of text read was longer in the former study. If



children had been given a longer passage to read, their reading speed might have been lower under parsed format conditions than under the standard format.

### *Discussion*

The error analysis determined that formats which simplify the syntax can improve reading comprehension for low ability readers but not for better readers. Low ability readers' error rates are reduced under both short sentence and parsed sentence formats. The source of the improvement appears to be explained partially by the amount of time spent on the passage. Although subject to individual variation, most low ability readers spend more time than do better readers on the nonstandard format text but the same time as others on regular text. This suggests that low ability readers ignore the greater complexity of or do not parse properly the more complicated intrasentence units. This supposition is supported by Weinstein and Rabinovitch (1971) who found that, in contrast to high ability readers, low ability readers do not improve in the number of nonsense words learned when syntactic cues are provided. The interpretation is further supported by the effects of the experimental formats and the correlations with reading ability (Figure 1 and Table 4). When intrasentence units are set out on separate lines or placed in separate sentences, low ability readers spend appropriately more time on the task and obtain comprehension scores that are closer to those obtained by better readers. High and middle ability readers do not spend more time on the experimental formats and are not helped by manipulated text. This suggests that better readers understand and locate while reading appropriate intrasentence units of text materials used here. It is possible that with text structures that are more complex than these, better readers could be helped by a parsed format; however, this remains to be shown.

Comprehension errors are not distributed in a linear fashion even when sentence length is held constant and word difficulty is increased systematically. The unexpected ease of the fourth section is interesting, particularly in the light of assertions that word length and sentence length are sufficient variables to measure text difficulty (according to commonly used readability formulas). One explanation of the low error rate on the passage section which contains the longer words (30 syllables per sentence) is that variations in text cohesiveness affect text difficulty (Stein & Glenn, 1977). The second and third sections consist of poorly related pieces of information (list-like structures) while the first and last sections have story-like structures. Since the comprehension question on all sections are of a factual nature and thus are of approximately equal difficulty, it can be concluded that scores on multiple choice tests of comprehension are affected by text cohesiveness as well as by word length.

### *Summary*

The first experiment demonstrates that children are not able typically to distinguish meaningful intrasentence units, particularly as sentences increase in length and contain longer words. Intrasentence junctures are inadequately identified, even by sixth graders, when compared with adults' judgments.

The second experiment indicates that only low ability readers' comprehension scores are improved when intrasentence units are separated. It occurs in conjunction with a reduced reading rate, perhaps because, under conditions when the syntax is more apparent, the low ability readers now process the more complex

information. It is conceivable, then, that separating sentences into parsed structures may help low ability readers cope with the syntactically more complex text materials that prevail from fourth grade onward.

Several questions remain unanswered. Will the effect be stronger: When longer sentences are given? When time constraints are imposed? If longer passages are used? If other measures of reading comprehension are given? Obviously there are a number of studies that should be undertaken before effects on comprehension of segmented sentences are fully understood.

Additionally, the fact that children do not agree with adults' conceptions of pausal junctures (Experiment 1) means that they might be trained to locate more of the necessary intrasentence boundaries. It is possible that this kind of task could improve comprehension of complex sentences. This effect needs to be demonstrated by a training study.

Nonconfirmation of a linear fit between comprehension error rate and word length is also important. Text complexity is presently based on only two factors, word length and sentence length. It is apparent that other variables such as story cohesiveness enter into text difficulty. As story grammars are more fully developed, the role in comprehension of text-level structures can be explored. This study suggests that text cohesiveness is an important third variable in assessing text readability.

#### References

- Buswell, G. An experimental study of the eye-voice span in reading. *Supplementary Educational Monographs #17*. Chicago: University of Chicago Press, 1920.
- Coleman, E., & Kim, I. Comparison of several styles of typography in English. *Journal of Applied Psychology*, 1961, 45, 262-267.
- Cromer, W. The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 1970, 61, 471-483.
- Eagan, R. *An investigation into the relationship of the parsing phenomena in oral reading and reading comprehension*. Unpublished doctoral dissertation, University of Alberta, 1973.
- Frase, L., & Schwartz, B. *Typographical cues that facilitate comprehension*. Murray Hill, N.J.: Bell Laboratories, 1977.
- Hartley, J., & Brunhill, P. Explorations in space: A critique of the typography of BPS publications. *Bulletin of the British Psychological Society*, 1976, 29, 97-107.
- Johnson, R. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 12-20.
- Levin, H., & Kaplan, E. Grammatical structure and reading. In H. Levin and J. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.
- North, A., & Jenkins, L. Reading speed and comprehension as a function of typography. *Journal of Applied Psychology*, 1951, 35, 225-228.
- Oakan, R., Wiener, M., & Cromer, W. Identification, organization, and reading comprehension for good and poor readers. *Journal of Educational Psychology*, 1971, 62, 71-78.
- Rode, S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*, 1974-75, 10, 124-141.
- Stein, N., & Glenn, C. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multidisciplinary approaches to discourse comprehension*. Hillsdale, N.J.: Ablex, Inc., 1977.
- Steiner, R., Wiener, M., & Cromer, W. Comprehension training and identification for poor and good readers. *Journal of Educational Psychology*, 1971, 62, 506-513.
- Weinstein, R., & Rabinovitch, M. Sentence structure and retention in good and poor readers. *Journal of Educational Psychology*, 1971, 62, 25-30.



MARILYN ASSHETON-SMITH

*The University of Alberta*

and

KELLEEN TOOHEY

*Ontario Institute for Studies in Education*

## School-Centered Community Conflict: The Holdeman Mennonite Case in Alberta\*

*The Holdeman Mennonites, or Church of God in Christ Mennonite, in 1976 began to establish private schools in eight rural communities across Alberta. This decision generated varying degrees of resistance from other community members, and varying amounts of conflict. This paper provides background information on the dispute, documents the conflict in two affected farming communities, and seeks to explain the difference in conflict through analysis of ideological and structural differences in the two communities. (Ms. Assheton-Smith is an Assistant Professor and Coordinator of the Intercultural Program in the Department of Educational Foundations at the University of Alberta; Ms. Toohey is presently studying towards her Ph.D. in Education at the Ontario Institute for Studies in Education.)*

The school is frequently the focus of community-based conflict. This conflict is often related to social divisions within the community, but the existence of social cleavage does not always generate conflict. Theoretical explanations for the presence of conflict in some cases and its absence in others continue to be only partially satisfactory. (See, for example, Gamson, 1966; Lee & Lapointe, 1977).

In Alberta, in 1977-78, a case of school-centered conflict occurred which represented a "natural experiment" for the study of community conflict, since in two closely comparable communities the same decision generated high levels of conflict in one and low levels of conflict in the other. The case involved the decision of Holdeman Mennonite parents across Alberta to withdraw their children from the local public schools and to establish independent Holdeman-controlled schools.

---

\*The authors would like to acknowledge gratefully the assistance of the Church of God Mennonite parents, school board members in both Camrose County and Spirit River School Division, as well as past and present members of the public school board and superintendents in both jurisdictions. The research was partially supported by an Operating Research Grant from the General Research Fund of the University of Alberta.



This study will first document the situation in the two communities, demonstrating thereby that the level of school-centered conflict was indeed high in one community and low in the other. It will secondly seek to explain why there might have been more conflict created by the Holdeman decision in one community than in the other. For explanatory purposes, two theoretical orientations will be drawn upon.

The first theoretical orientation is derived from Carlton (1974), who states that different images of the school held by different sectors of the public may explain school-centered conflict in a given situation. Carlton argues that "an understanding of the way in which people think about the school is relevant to their behaviour in relation to the school" (1974, p. 73) and that at times some images of schooling have achieved very wide acceptance, becoming important influences in the development of school systems.

The description of a relationship between thoughts, or images of schools, and actions implies a kind of idea system which Guy Rocher would term an ideology. These ideas have, in Rocher's terms, a conative function: "pushing . . . a collectivity to action, or at least directing the collectivity by providing goals and means" (1972, p. 103). Translating Carlton's statements into Rocher's terms,<sup>1</sup> school-centered conflict at the ideological level would be based upon different ideologies of schooling held by different groups.<sup>2</sup>

The second theoretical orientation is generated by adding to Carlton's notions a perspective provided by Gerth and Mills (1953)<sup>3</sup> and the general sociological orientation given by Coulson and Riddel (1970). Carlton views his school images, or school-related ideologies, as dependent upon the social institution to which the school is believed to be responsible. These include, among others, the family, the state, religious groups, and professional teachers. Gerth and Mills (1953, p. 251), dealing with a very similar set of social institutions, would convert Carlton's images to theoretical types with the argument that schools have little autonomy from other social institutions. That is, it is not just a matter of beliefs, because schools are simply extensions of one or another of the social institutions named by Carlton. If this is the case, conflict around schools should not be conceptualized merely as a conflict of "images," but rather as conflict relative to which of the other social institutions will, in fact, control schooling and thereby determine the nature of schools.

Coulson and Riddel potentially extend this type of conflict analysis by treating social analysis as the process of analyzing groups in interaction within a social structure which benefits some groups more than others. Conflict arises when one group seeks to improve the benefits to itself and another group recognizes this as a threat. The details of the conflict depend upon the consciousness or understanding of the situation which each of the participant groups have. But the analyst must seek to understand both the manifest explanations of the conflict, which are ideologically related explanations given by the participants themselves, and the latent explanations which are the causes embedded in the social structure. The actors themselves may not be aware of these structural causes.

#### *Methodology and Data Collection*

The methodology of this research consists of a comparative case study approach, with two cases selected on the basis of variation on one dependent variable: the level of conflict created by the Holdemans' decision to establish

independent schools. Both communities are described structurally and documentation of the school-related beliefs (or ideologies) of Mennonites, professional teachers, and school board members of Camrose County and Spirit River Division is presented. The authors then consider whether ideological differences or structural differences account more adequately for the observed differences in the dependent variable.

The data collection procedures were two-fold. First, the research team conducted unstructured interviews with persons who were members of each of the categories outlined above (Mennonites, professional teachers, and school board members). These were carried out in the homes or offices of the respondents. The interviewers sought both to obtain information about the events surrounding the establishment of the independent schools and to obtain information on the beliefs about schooling held by the respondents. The second data collection procedure consisted of the examination of available data from county and school division offices, and from the Department of Education in Edmonton. These included documents such as annual reports, enrollment records, financial records, and county maps.

Neither the cases selected nor the respondents interviewed should be considered representative as that term is used in probability sampling. Rather, all were selected with the above kinds of theoretical problems in mind. At present, inferences must be drawn with caution because of the exploratory nature of the research, and generalization to other similar cases should also be made with care.

#### *Background Information: The Holdemans and Alberta Schools*

The Holdeman Mennonites are one of a number of subdivisions of the followers of Menno Simons. Their historical origins were in sixteenth century religious conflict in Central Europe and involved both a rejection of the Roman Catholic authority and rejection of involvement with the civil authority. They could be seen then as early proponents of the separation of church and state, but also as early practitioners of adult-shared community decision-making, based upon biblical teachings rather than on response to an hierarchical authority structure.

A number of Mennonites migrated to Canada beginning in 1874 because of the availability of land in Manitoba and the assurances of the Canadian government that they would have freedom from military service and the right to their own schools. One Joseph Holdeman, himself originating from a different migration of Mennonites to the United States, began preaching a conservative revivalism in 1858; the Alberta Holdeman group owe their intellectual origin and contemporary conference membership to this branch of the Mennonite faith.

The Holdemans began their settlement in Alberta at Linden, near Three Hills, in 1902. In the 1930s, some families headed north to settle at Crooked Creek in the Peace River country and in the early 1960s, families began to move into the Spirit River area and into Camrose County. Congregations currently exist in these as well as in the Stettler and Pincher Creek districts. The total number of persons involved is approximately 800, representing a small fraction of the total Mennonite population in the province of Alberta.

The Holdemans began to establish their private schools in Alberta in 1976. At the time of the research, the two Holdeman schools in Spirit River School Division were in their second year of operation and those in Camrose County in their first



year. By 1978, there had been a total of eight such schools established in the province with, at most, 300 children in attendance.

The 1978 court case in Alberta involved essentially an attempt to return Holdeman children to the public school at Linden (near Three Hills) through legal intervention. Charges were laid against forty-five Holdeman parents (with one parent being the test defendant) because their children were truant from a school approved by the Minister of Education. The Holdeman Schools were not approved because they did not have certified teachers and had not been inspected by school authorities. In a decision handed down on February 6, 1978, Judge Oliver ruled that the test defendant was not guilty, by virtue of the argument (among others) that specifically, in the case of the Holdemans, the School Act conflicted with Section 2 of the Alberta Bill of Rights guaranteeing religious freedom. During the first week of March, 1978, Julian Koziak, the Minister of Education, announced that a fourth type of approved school would be added to the Alberta School Act. This type of school could operate without certified teachers, but would receive no funding from the government. This decision removed, at least for the time being, the legal threat to the Holdeman schools.

### *Description of the Conflict*

As noted earlier, two of the Alberta school jurisdictions which experienced Mennonite withdrawal of students seemed to respond quite differently. In Spirit River School Division, although board members were concerned about the Mennonite decision and participated in meetings with other boards and the Department of Education to discuss the situation, they stated publicly that the Holdeman Mennonite schools seemed adequate. In Camrose County, fear of the school system being destroyed was expressed and parents petitioned against the Mennonite school being allowed to continue.

At the local community level, the difference in intensity of response or reaction was most marked. At Hillstream in Camrose County, the Home and School Association circulated a petition against Mennonite schools in February of 1977 and, in a matter of days, had gathered sixty signatures. They presented their petition to the County School Board, asking members to read and to sign it. The School Board refused to sign the petition as a Board but agreed that board members could sign it as individuals (*Camrose Canadian*, February 2, 1977, p. 1).

The Home and School meeting in Hillstream in April of 1978 had as its topic, the Mennonite school situation, and as its guests, representatives of the provincial Department of Education and the Alberta Teachers' Association. Fifty persons attended that meeting, a fact which suggests that neither the court decision nor the minister's legislation satisfied community objections to Holdeman schools. By contrast, there was no organized reaction in the Spirit River district.

There was also considerable difference at the local level in the newspaper coverage of the topic. Residents of the Spirit River School Division are served by a number of regional newspapers. As far as could be ascertained, there was no newspaper coverage of the withdrawal of Holdeman Mennonite children from the area's public schools. A regional paper, the *Grande Prairie Herald-Tribune*, began in September of 1977 to carry Canadian Press stories of the court case at Linden but contained no investigation of, or even reference to, the local private Mennonite schools. The smaller papers as well printed no stories of the local issue, nor did



they carry the Canadian Press stories about the court case at Linden. In fact, very few stories relating to schools were carried in these papers.

On the other hand, the regional newspaper in Camrose, the *Camrose Canadian*, routinely reports on County council meetings, and thus on school business. As early as February, 1977, the *Camrose Canadian* was reporting on school board discussion of the Mennonite proposal to withdraw their children. It also printed letters to the editor on the subject and generally covered news of the conflict.

### *Description of the Two School Jurisdictions*

#### *Case A: Camrose County and "Hillstream" School*

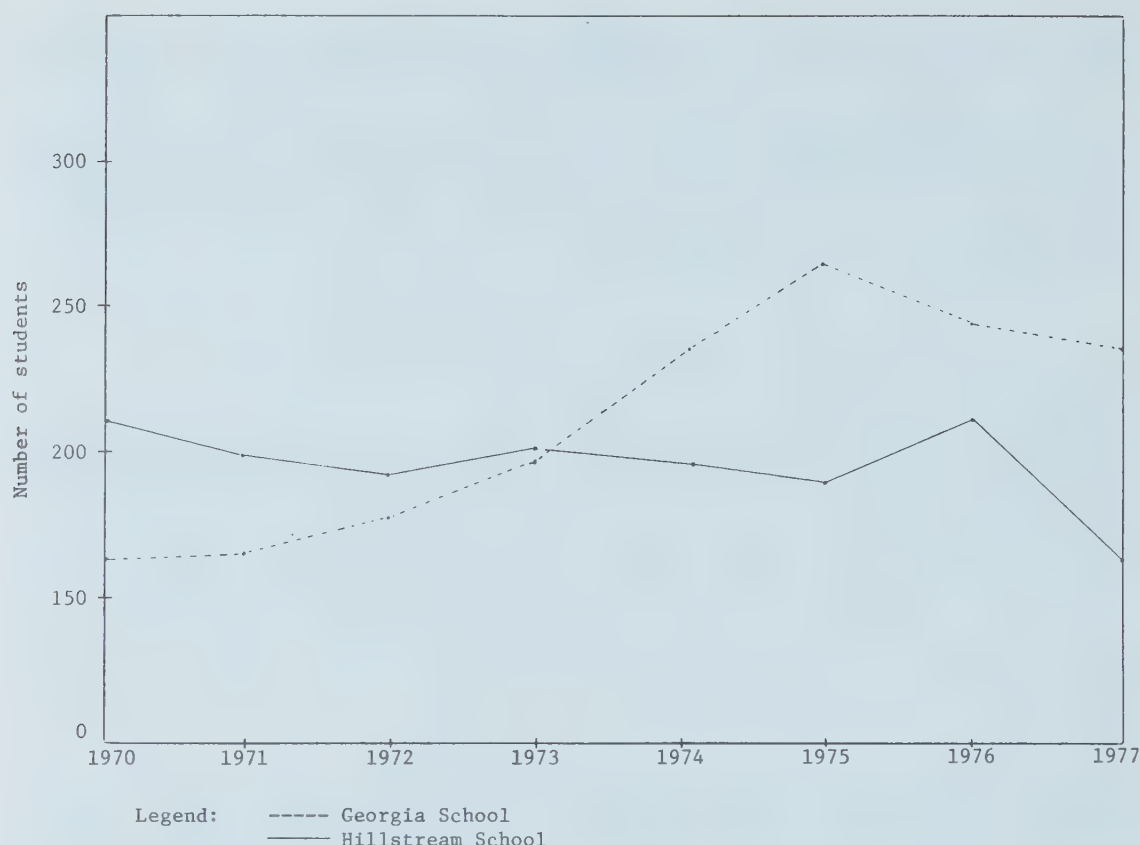
Camrose County is located in central Alberta, some 100 km southeast of Edmonton. County offices are located in Camrose, a city with a population of 11,000, which has its own municipal government and school boards. In addition to the 3,500 square km of farm land under county jurisdiction, there are eight incorporated villages and towns, the schools of which are part of the county system.

School centralization resulted in all country schools being centralized into villages and towns. There is no separate school within the county and, until the Holdeman Mennonite decision, there were no private schools. However, county students may enroll in the Camrose Separate (Roman Catholic) school, or in Camrose City high schools. In 1977, 2,136 students were enrolled in the county system, and 264 county residents were enrolled in other school systems. To these 264 must be added the 55 Mennonite children; that is, 17% of all resident children enrolled outside the county schools are Mennonite. All of these 319 children represent a per capita funding loss to the county in relation to provincial funding of schools.

Hillstream school, from which the Holdemans withdrew their children, is located on the edge of Hillstream village, whose population is about 140; no Holdeman Mennonite families reside in the village. The farm population began to settle in the area on homesteads in 1899; Mennonite families started to purchase farms in 1959. Mennonites now make up about 25% of the ninety farm families in the district, but their children represent 50% of the farm school children because of their larger family size. School enrollment was just over 200 in 1962, declined to 160 in 1968, but jumped to 210 in 1972 with the addition of grades nine to twelve from a nearby village. Enrollment remained fairly stable (near 200) from that time until the withdrawal of the Holdeman children in the fall of 1977 when enrollment dropped to 155, with grades one to nine dropping from 143 to 99 students (see Figure 1).

The number of teachers on staff reached its maximum of 14 in 1976; it was down to 12 in 1978. With five of the teachers working at the high school level, there is some doubling of grades at the elementary level. Of the teachers on staff, half grew up in the county and generally attended schools much like Hillstream, and four were raised in the immediate area. The majority of teachers live in the village or on nearby farms, and one is presently mayor of the village.

The Holdemans established their schools in the fall of 1977. There are two schools, one at the southeast edge of the school enrollment area and one in the northwest corner. In each case, the teacher is a man schooled to about the grade ten level in public school. Also teaching or assisting at each school are women who have completed grade ten or eleven. The schools teach from grades one to nine.



*Note.* Holdeman children were withdrawn from Georgia School in 1976, from Hillstream School in 1977.

Figure 1. School enrollment, 1970-1977: Hillstream and Georgia schools. (Source: Annual Reports, Camrose County No. 22, 1970-1977; Spirit River School District No. 47, 1970-1977.)

### *Case B: Spirit River School Division and "Georgia" School*

The municipal level of government in this case consists of Municipal Districts and Improvement Districts. The School Division, a separately elected and funded structure, incorporates all of the territory in the Municipal District of Spirit River plus much of the land area of two Improvement Districts. It is located in the Peace River area, some 600 km northwest of Edmonton. The School Division offices are located in the town of Spirit River which has a population of just over 1,000. This town's public schools are within the School Division. There are six incorporated towns and villages in the Spirit River School Division, but ten schools, because some of the schools were centralized in villages and towns whereas others were centralized in farming districts without associated villages.

The total enrollment in the school system in September 1977 was 1,686 students, but an additional 170 resident students attended other schools (Roman Catholic Separate School; Seventh Day Adventist and Mennonite Private Schools, and a nondivisional high school). Of these 170 students, just 28, or 17%, are Holdeman Mennonites — exactly the same percentage as in Camrose County.

The Georgia farming community, unlike Hillstream, has no village within its boundaries, although a store is situated across from the school and a recreation complex was opened recently. These three buildings — the school, store, and

recreation complex — provide a minimal service centre for the surrounding area and as such are roughly equivalent to the village in the previous case. However, the majority of the population of this centre are teachers living in eight mobile homes on the school grounds. The farm community was opened to homesteading in the 1950s, at which time Holdeman Mennonite farmers and others began to move in, clear the land, and settle. There are fewer than twenty Holdeman families in the area, representing less than 10% of the total farm population.

Georgia school was first established as a three-room school in 1965; rooms have been added on four occasions since that date. In 1976, a high school addition was constructed, enabling the first children to start grade one in Georgia school to complete their grade twelve there. School enrollment has shown a general increase, from 60 in 1965 to 230 in 1978. The Holdeman Mennonites withdrew their children from the school in 1976, resulting in an enrollment loss of only 13 children (see Figure 1).

The highest complement of teachers in this school was 15 in 1976, dropping to 13 in 1977 and remaining the same in 1978. Of those teachers, only one is from the Peace River country, the rest being from outside the province and in many cases from outside the country. In the past year or so, there has been some sign of this teacher population stabilizing, with about half purchasing land in the area or building homes and apparently deciding to settle. As such, their recent and diverse origins are not unlike those of their neighbours who also have moved into the area within the last ten or fifteen years.

The Holdeman Mennonite school has one room and offers grades one to nine to thirteen students.

#### *Data on Group Beliefs*

The following material represents primarily the data collected through interviews dealing specifically with each group's "school-ideologies" or beliefs about schools and their understanding of the situation about Holdeman withdrawal of students from public schools. In some cases, this has included beliefs about the society and about the actions or beliefs of other groups involved. The data were collected between March 10 and April 10, 1978, prior to a media focus in Alberta on an incident in the Mennonite school at Ridge Valley. Such a media event could potentially have modified some of the responses reported here.

#### *Holdeman Mennonite Beliefs*

The Holdeman Mennonites' statements about their beliefs and understanding of the situation were consistent in both communities investigated. Although they did not necessarily object in major ways to the nearby public schools which their children had attended, they did object to the general institution of public schooling. Society is perceived as losing its moral focus and this is mirrored in the public schools. At one time rural schools were somewhat protected from events in urban areas, but this is no longer felt to be the case and the rural schools now have many of the problems of urban schools. The "worldly" task of the Mennonite is to follow scripture; this task is felt to be best carried out on the economic base of farming. Parents and the community cannot ignore their responsibility to guide their children into adult roles.

In contrast to the public schools, Holdeman schools see as their goal teaching children the 5 R's — reading, writing, arithmetic, responsibility, and respect.



School is a place of work; recess is a time for play. Adults who have a proper relationship with God and who are committed to teaching can teach the desired skills and virtues. In fact, the most important criteria for choosing teachers is spiritual righteousness and dedication to teaching. It is felt that children can complete the standard curriculum in eight months, but as there is no desire to create undue tension with neighbours, they are kept in school for almost the standard school year. The Holdeman curriculum follows Alberta requirements in general, although texts with appropriate biblically based discussion are used. The school board members of the eight established Holdeman Mennonite schools in Alberta assist each other in the selection of curriculum material. In Spirit River School Division, it was stated, the public school donated old texts for use in the new private schools.

The need to maintain autonomy from the state was clearly indicated. As one respondent noted, "In Manitoba first they provided the firewood, then they paid for the teachers, then they built the school . . . then they started to tell us how to run it." The Holdemans expressed little interest in the working of government; they pay their taxes but do not attempt to determine where the money goes or how it is used. The above Holdeman belief system was consistent across both school jurisdictions and thus had the nature of a "shared ideology," a well-articulated, systematized belief structure which forms the basis for action or legitimizes action once it is taken.

### *Professional Teacher Beliefs*

Teachers in both jurisdictions were fairly consistent in their beliefs. As a result of the situation, they tended to initiate discussion about school beliefs by emphasizing the necessity of teacher certification for the teaching task. Although it was stated that some persons might be able to teach without university training, the general notion was that certification provided a control of standards and expertise that would ensure the best teaching.

There was also a belief expressed that teacher certification and the structure of supervision associated with public schools would result in the actual following of the Alberta curriculum in the classroom. It would seem that belief related to both the teacher-learned skills associated with curriculum study and the notion that teachers accept the legitimacy of the Alberta school curriculum so they would follow it even if they were not under supervision. Teachers expressed the belief that this curriculum ensures a certain minimum level of education in schools. The extent to which they expected Mennonites to deviate from that course of study was a real concern. They expressed particular concern that science would not be taught in the private schools.

Teacher respondents did not always find the goals of schooling easy to articulate. Although there were references to the need to learn the 3 R's, the desired consequences of schooling expressed were a great deal more complex: developing social skills, acquiring a level of competence which enables students to advance to other educational institutions, being able to make a living and contribute both to society and to the employment sphere.

The only clearly expressed state-related role described by teachers was enforcement of the law by the courts: if the law demands compulsory school attendance, certified teachers in the school, or approval of private schools, the courts should enforce those laws. The general impression of teachers was that the

courts tended not to enforce these laws, even in cases other than the Holdeman dispute. Teachers referred to the problems of spending four years obtaining a university degree, becoming a certified teacher, and then having employment prospects threatened by the development of schools which would use uncertified teachers. The belief was expressed that the school boards would now find ways to use uncertified teachers as less costly personnel, and that eventually independent schools such as those of the Mennonites would seek and obtain government funding.

The above described teacher belief system was consistent across both jurisdictions and like that of the Holdemans, it could be seen as part of a “shared ideology.”

### *School Board Member Beliefs*

All interviewed school board members from both jurisdictions expressed a general feeling of discontent with the public schools as they are, echoing some of the reasons the Mennonites gave for disapproval of public schools in general; in fact, almost all initiated discussion by indicating agreement in some sense with the Mennonites’ analysis and subsequent decision to withdraw their children.

School board members saw teacher certification as an issue of importance primarily to teachers, but were very concerned with the cost of schooling. In some cases, this meant objecting to “costly frills” such as outdoor education; in other cases, it meant objecting to the cost of certified teachers and the teachers’ union.

The source of the problem in this economic squeeze was felt by board members of both school jurisdictions to be related to inadequate funding. In Camrose County, school board members were more likely to be struggling with the notion of reducing costs whereas in Spirit River, they were attempting to persuade the government to recognize the inadequacy of current funding policy for their particular situation.

In 1975, there was a major funding crisis in both school jurisdictions. Both made an attempt to pass bylaws permitting a funding increase, bylaws defeated by the electorate in both cases. In Camrose, this was used as justification to close two schools, but no schools were closed in the Spirit River School Division.

### *Analysis and Conclusions*

Following the interviews and extensive search of local press coverage in each community, it was evident that there actually was a higher level of conflict and resistance to the Mennonite decision in the one community than in the other. However, it was apparent that there was more conflict at the local community level than at the larger school jurisdiction level. In the one case, an active Home and School Association provided a structure for expressing resistance while in the other case, there was no such structure nor evidence of resistance. At the county or school division level, the major and, at times, desperate problems of school financing seemed to override any emphasis on the Mennonite withdrawal itself. This was the case particularly in Spirit River.

Why would there be such a difference in the degree of conflict between the two cases? Was it strictly ideological, a conflict between opposing belief systems, or were there significant structural differences between the two cases?

At the ideological level, the Holdeman Mennonite beliefs were totally consistent



in both communities, and consistent with news reports of their beliefs and motivations elsewhere in the province. This consistency tended to be interpreted by others as evidence of obedience to an authority structure but it is also consistent with a fully articulated shared belief system which is held by Holdemans in Alberta and elsewhere. Provincial, national, and international conferences (such as the one in Kansas in 1973 which apparently sparked the schools' decision) would provide settings for exchange of information, world-views or ideologies, and exchange of proposals for action appropriate to the beliefs.

Similarly, the teachers were, on the whole, consistent in their presentation of a professional ideology. Their belief system includes not only the notion of teaching expertise and standards linked to certification, but also some ideas about curriculum which, they appear to believe, should be scientific. They expressed opposition to the "religious" curriculum they believed the Mennonite schools would use, but generally believed that the public schools could accommodate the Mennonites and that it should not have been necessary for the Holdemans to withdraw their children. This position was expressed more strongly in Camrose County than in Spirit River. In the case of Hillstream, whereas withdrawal of the Mennonite children was acceptable in principle as a religious or parental right, it was not acceptable in practice, as it threatened the ongoing public school system.

Members of school boards showed less ideological consistency. Persons interviewed were concerned about what was being learned in schools. They were particularly worried about the cost of schooling. In Spirit River, this concern was partially related to excessive school expenses, though the main focus was on government and lack of adequate government funding. It would seem that in Camrose this emphasis was reversed.

There were no significant differences in ideological orientation between the categories of respondents from the two communities. Ideology, as such, would therefore not seem to explain the differences in conflict in the two cases. What differences, then, exist at the structural level which might be considered causal of the different levels of conflict? There appear to be three of particular significance.

In the first place, the Hillstream school is much more seriously threatened by the withdrawal of Mennonite children. It has lost 25% of its school enrollment, while Georgia school has lost only 5%. Hillstream's teacher complement is reduced by two, a saving at the county level which Spirit River might envy (Spirit River was not able to make reductions in teaching staff), but which is certainly a threat at the local level. Whereas Hillstream may see a smaller and smaller school in the future, Georgia school continues to expand, with grade twelve having been added in 1977. The first case is one of retrenchment, the second of expansion.

It may be the case that this first factor, a structural factor associated strictly with schooling itself, would have been enough to generate the differences in levels of conflict. Yet further structural aspects must be noted. The Hillstream residents and district farmers see a threat not only to their school but to their whole community. A viable, active community of farmers, a relatively homogeneous community which lacked even the pluralism of separate schools, much less private schools, has had 25% of its population replaced by Holdeman Mennonite farmers. Unlike in the Spirit River district where Holdemans and others moved in at the same time, in Hillstream the Mennonites may be seen as "invaders" in the sense that term is used in ecological analysis of the Chicago "school of sociology."



Finally, in the Hillstream case the school teachers are an integral part of the community. As noted earlier, one is mayor of the local village, others own homes and farms in the area, and about half grew up in the locality. The community and the teachers can act relatively as one: the teachers' ideology may motivate their actions but their shared social relations with kin and neighbours partially enables them to transmit their ideology and to act as motivators in the community's organized resistance. By contrast, the teachers at Georgia are outsiders; living in mobile homes on the school grounds dramatically symbolizes this relationship. Some are beginning to farm in the area, and eventually school and community will no doubt become more articulated, but at present, neither the teachers' ideology of education nor their problems are shared through well-developed social relations.

There are, no doubt, other factors; other communities could be examined to determine if all of these elements are present. The role of the media in clarifying an ideological situation should be examined somewhat more closely, as well as the significance of an organized structure for parents to discuss their school concerns with school staff. However, the three components described above — structural threats to the school, a replacement threat to the community itself, and ideological conflict spread by one interest group to a larger community through shared social relations — are present in the situation where conflict is high and absent in the situation where conflict is low.

There are a myriad of other possible inferences and directions for further research suggested by this study. For example, it was stated earlier in the paper that "conflict around the school should not be conceptualized merely as a conflict of 'images' but rather as a conflict as to which of the other social institutions will, in fact, control schooling and thereby determine the nature of school" (Supra, p. 78). It is clear that the Holdemans have established their schools as extensions of the family and of religion, and that they very explicitly see control as appropriately and necessarily resting with these two institutions. Although they will accept guidance in relation to curriculum, they explicitly reject any assumption of the right of the state or state-certified teachers to exert control over these schools. Have they, then, created a school which in theoretical terms is a "type" of institution different from the public school? Would the literature on school effects,<sup>4</sup> or consequence of schooling, apply to this type of school?

Professional teachers, on the other hand, do not see the appropriate authority for control of schooling to be the family or a religious institution. They see schooling control to be the province of those trained and licensed to perform that function. The degree to which teachers admit the legitimacy of the state to exert control is indeterminate: in terms of teacher certification, enforcement of legislation, or setting of official curriculum, teachers seem to accept state control because it is supportive of professional teacher interests and belief systems. As the Minister of Education's decision to accept a fourth category of approved private schools (one that does not require certified teachers) is not supportive of teacher interests, they are confused by it. In Rocher's terms, they are unable to give meaning to the Minister's actions which cannot be interpreted within the professional ideology.

Finally, one of the many omissions from the study is an analysis of the actions at the "centre," at the level of the government. By focusing research at the community level, we have no information on the beliefs or ideology and structural aspects influencing the decisions of the Attorney General's department, the

Minister of Education, the Cabinet, the Department of Education, or the Alberta Teachers' Association.

Notes

1. Rocher's (1972, p. 103) particular use of the term "ideology" needs amplification. He considers the term ideology to be close to W. I. Thomas' "definition of the situation," which is the way that members of a collectivity explain their present situation and give it meaning. However, he adds three additional points: (a) ideology has a systematic, coherent and organized form because it is explicit and verbalized; (b) ideology refers a great deal to values; (c) ideology has a conative function — it pushes or incites a collectivity to action, or at least directs this collectivity by providing goals and means.
2. The phrase "hegemony of beliefs" refers to an idea system which is so dominant or widespread that there is little or no conflict in a society. Thus, the alternative to Carlton's ideological conflict would be ideological hegemony.
3. We are indebted to our colleague, Professor Urion, for his observations on this point.
4. Alan Pomfret implicitly addresses a similar question in his 1977 paper on Newfoundland schooling. He basically argues that the structure of Newfoundland schooling reflects the "periphery" status of Newfoundland. This paper, partly a study of denominational schooling in Alberta, challenges some of his conclusions.

References

- Alberta, Bill of Rights, 1972.
- Archer, W. *Education and the Mennonites*. Unpublished paper, University of Alberta, February, 1978.
- Camrose Canadian*. Camrose, Alberta: February 1977 to February 1978.
- Carlton, R. A. Popular images of the school. *Alberta Journal of Educational Research*, 1974, 20, 59-74.
- Coulson, M., & Riddell, C. *Approaching sociology: A critical introduction*. London: Routledge and Kegan Paul, 1970.
- Epp, F. H. *Mennonites in Canada 1786-1920: The history of a separate people*. Toronto: Macmillan of Canada, 1974.
- Gamson, W. Rancorous conflict in community politics. *American Sociological Review*, 1966, 31 (1), 71-81.
- Gerth, H., & Wright Mills, C. *Character and social structure: The psychology of social institutions*. New York: Harcourt, Brace and World, Inc., 1953.
- Klassen, P. G. *A history of Mennonite education in Canada 1786-1960*. Thesis, University of Toronto, 1970.
- Lee, D., & Lapointe, J. Conflict over schools in a multi-ethnic society. In R. A. Carlton et al. (Eds.), *Education, change, and society*. Toronto, Gage, 1977.
- Pomfret, A. Historical change in a peripheral society: A working paper on Newfoundland. Paper presented at the meeting of the Canadian Sociology and Anthropology Association, Fredericton, New Brunswick, June, 1977.
- Rocher, G. *A general introduction to sociology: A theoretical perspective*. Toronto: Macmillan of Canada Ltd., 1972.

DAVID M. WILLIAMS

*The University of British Columbia*

## A Study of Moral Education in Surrey, B.C. Secondary Schools

*This paper reports the assumptions, procedures, findings and implications of the second year of a two-year study of moral education. Year 2 of the study involved 411 junior and senior secondary school students and 13 teachers. Groups in two experimental moral education conditions were compared with each other and with a control condition on a battery of measures of moral reasoning abilities and attitudes. The main findings were of significant treatment effects on normative reasoning and nonprejudice measures attributable to one of the moral education conditions. Hypotheses concerning treatment group differences in distinguishing factual and value claims and moral development were not confirmed. Implications of discrepant findings and procedural problems for further research and development are discussed. (Dr. Williams is an Associate Professor in the Social Studies Department, Faculty of Education, the University of British Columbia.)*

Workers in the field of moral education and research face a range of problems, including (a) lack of agreement concerning the role of schools in promoting moral development, (b) lack of convincing theory concerning the requirements of mature moral judgment and action, (c) lack of effective means for the measurement and/or promotion of moral components, and (d) limited knowledge of the requirements for training or selecting teachers for work in moral education. Although problems of theory, test and curriculum development, and in-service education beset all areas of public education, they are particularly critical in a problematic and controversial field such as moral education.

The Association for Values Education and Research (AVER, Faculty of Education, U.B.C.) has conducted a multidimensional research and development project in cooperation with junior and senior secondary schools in Surrey, B.C. The project has been sufficiently comprehensive and flexible to allow for experimental comparisons of alternative approaches to moral development, as well as concurrent development work in curriculum, measurement, and teacher training.

The activities, problems, and findings of Year 1 were reported in the 1975 Yearbook of The Canadian Society for the Study of Education (Williams, 1975).



The purpose of the present paper is to describe the assumptions, activities, findings and implications of Year 2 of the AVER-Surrey Project.

*Assumptions Underlying the AVER-Surrey Project*

AVER's work is based on the assumption that any adequate account of the requirements of mature moral judgment and action must be at least as comprehensive as that shown in Figure 1. Those of you who are familiar with the moral education literature will recognize the influence of John Wilson (1972), Coombs and Meux (1971), and earlier AVER research (AVER, 1974, 1975) on this "working list" of moral components. AVER assumes, in addition, that there are developmental constraints on the acquisition of moral reason-giving (Kohlberg, 1971), social perspective or role-taking (Selman, 1975), and other moral components.

- 
1. Disposition to regard others as equals and to give weight to one's own and others' interests (CONCERN FOR OTHER PEOPLE).
  2. Ability to know what others (and yourself) are feeling and what their (and your) interests are (EMPATHY/ROLE-TAKING).
  3. Knowledge of facts relevant to moral choices (SOCIAL KNOWLEDGE).
  4. The 'know-how' to act effectively in social contexts (SOCIAL SKILLS).
  5. Ability and disposition to give reasons for moral judgments and actions *and* to assess reasons given in terms of factual warrant *and* moral justification (REASON-GIVING, FACT ASSEMBLY and PRINCIPLE TESTING).
  6. Sensitivity to morally hazardous situations, including
    - a. actions that may have consequences for others that one could not accept for oneself, and
    - b. actions that may have disastrous consequences if everyone were to engage in them (MORAL POINT-OF-VIEW).
  7. Resolution to act on justifiable moral judgments: to do what one has decided is right and to refrain from doing what one has decided is wrong.
- 

Figure 1. Moral abilities, attainments, and dispositions.

These assumptions suggest important implications for moral education theory and research. They imply, for example, that it is unlikely that a single instructional strategy will promote all of the moral components. Another implication is that developmental constraints will limit the degree to which some, if not all, of the moral components will be amenable to instructional manipulation within limited treatment periods. Finally, these assumptions imply for AVER that the structural analyses of global stages being undertaken by researchers in the developmental psychological tradition, such as Kohlberg and Selman, can at best provide only a partial picture of moral development and education. The ability of researchers to weigh the relative importance of such components as those identified in Figure 1 in predicting mature moral judgment and action will depend in part upon advances in the measurement and promotion of discrete components. This moral components approach, rather than the developmental psychological orientation, is predominant in the AVER-Surrey Project.

Curriculum Development: Experimental Conditions

In Year 1 of the project, AVER developed curricula, trained teachers and evaluated the effects of three approaches to values and moral education with respect to each other and to a non-values, control curriculum. Junior secondary school students in grades eight, nine and ten were the subjects of the Year 1 study. As a result of Year 1 activities and findings, a second study involving further curriculum development, teacher training, test modification and comparison of approaches was undertaken in Year 2 of the project.

On the basis of Year 1 findings, it was decided to combine elements of the Meux (1974) programmed text materials (known in Year 1 as the Meux condition, or approach) with the reasoning elements of the AVER condition to comprise a modified AVER condition for implementation and evaluation in Year 2. The Kohlberg condition was retained in modified form in view of Year 1 and independent evidence (Kohlberg, 1971) suggesting the promise of this approach for the enhancement of some of the moral components. Unlike Year 2, funding did not permit curriculum development and teacher training for a non-values control condition. However, it *was* possible in one of the participating Year 2 schools to assign a proportion of grade 11 and 12 students to regular English and Social Studies classes, which comprised a quasi-control condition. In effect, therefore, Year 2 involved curriculum development and teacher training for two experimental conditions, AVER and Kohlberg. Testing was conducted with students under three conditions, AVER, Kohlberg, and control.

Experimental Condition 1. The AVER Approach

AVER believes that a morally autonomous, or morally educated, individual must bring to bear upon any moral problem a number of attainments, abilities and

Theme	Tasks
Unit 1: Status of Woman	Facts/Values Distinction; Clarifying Value Objects; Role Playing; Practical Syllogisms.
Unit 2: Population	Fact/Value Distinction, Identifying Point of View; Fact Assembly; Rudimentary Procedure.
Unit 3: Pollution	Fact/Value Distinction; Role Playing; Fact Assembly; Rudimentary Procedure; Principle Testing; Role Exchange Test.
Unit 4: Prejudice	Determining Relevant Differences; Fact Assembly Chart; Practical Syllogism; Principle Testing; Role Exchange Test; Universal Consequences Test; New Cases Test; Subsumption Test.
Unit 5: War and Revolution	Clarifying Value Objects; Point of View: Moral Viewpoint; Practical Syllogisms; Indicators of Relevance.
Unit 6: Old Age	Fact/Values Distinction; Point of View; Role Playing; Fact Assembly Chart; Practical Syllogisms; Principle Testing; Role Exchange Test; New Cases Test and Subsumption Test.
Unit 7: Capital Punishment	This was designed as the final unit in which all AVER Tasks were reviewed.

Figure 2. AVER tasks by theme and unit.

dispositions. Various exercises designed to foster the acquisition of moral reasoning and role-taking components were incorporated into the AVER condition curriculum materials. AVER exercises included: (a) distinguishing factual and value claims, (b) identifying point-of-view, (c) fact assembly and rudimentary procedure, (d) the practical syllogism, (e) principle testing, and (f) role playing.

AVER units focusing on seven moral value issues were assembled. The ideal of having each unit focus upon a particular ability as well as serve as a cumulative review of the other moral components was only partly achieved, because of limited time and resources. Figure 2 identifies the tasks emphasized in individual AVER units.

### *Experimental Condition 2. The Kohlberg Approach*

Lawrence Kohlberg (1971) claims on the basis of extensive longitudinal, cross-cultural and experimental research that moral reasoning develops through an invariant sequence of stages. He and Turiel (1973) have found that students can comprehend and tend to prefer reasoning at one stage higher than their own dominant stage. Furthermore, Kohlberg claims that movement towards the next

---

#### Level 1 — Premoral

Stage 1 — Orientation toward punishment and unquestioning deference to superior power. The physical consequences of action regardless of the human meaning or value determine its goodness or badness.

Stage 2 — Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others. Human relations are viewed in terms like those of the market-place. Reciprocity is a matter of "you scratch my back and I'll scratch yours", not of loyalty, gratitude or justice.

#### Level II — Conventional Role Conformity

Stage 3 — Good-boy-good-girl orientation. Good behaviour is that which pleases or helps others and is approved by them. There is much conformity to stereotypical images of what is majority or "natural" behaviour. One seeks approval by being "nice".

Stage 4 — Orientation toward authority, fixed rules and the social order. Right behaviour consists of doing one's duty, showing respect for authority and maintaining the given social order for its own sake. One earns respect for performing dutifully.

#### Level III — Postconventional Morality

Stage 5 — Social-contract orientation, generally with legalistic and utilitarian overtones. Right action tends to be defined in terms of general rights and in terms of standards which have been critically examined and agreed upon by the whole society. There is a clear awareness of the relativism of personal values and opinions and corresponding emphasis upon procedural rules for reaching consensus.

Stage 6 — Orientation toward the decisions of conscience and toward self-chosen ethical principles appealing to logical comprehensiveness, universality and consistency. They are not concrete moral rules like the Ten Commandments. Instead, they are universal principles of justice, of the reciprocity and equality of human rights, and of respect for the dignity of human beings as individual persons.

---

Figure 3. Kohlberg's stages of moral development.



higher stage of moral reasoning can be enhanced if subjects are given opportunities to hear “+1” reasoning during group discussions of difficult moral choices, or dilemmas. A brief description of Kohlberg’s six stages of moral development is given in Figure 3.

Curriculum materials for the Kohlberg condition focused upon the same unit themes as the AVER condition. However, the materials comprised Kohlbergian moral dilemmas relating to each of the unit themes. These were presented in print, sound filmstrip, and role playing formats. Questions accompanying the dilemmas encouraged pupils to make choices and to give reasons for those choices. Teachers were expected to promote reason-giving at a variety of stages of Kohlberg *and* to promote lively exchange of reason-giving on the part of pupils.

### *Teacher Training*

Although a total of 13 teachers participated in Year 2 of the AVER — Surrey Project, only the 11 teachers assigned to one or other of the experimental conditions took part in training sessions conducted prior to the implementation period. The two control condition teachers were not trained by AVER personnel. Control teachers implemented their regular English or Social Studies programmes. It must be acknowledged, therefore, that unlike Year 1, comparisons of the experimental conditions with the control condition fail to control for a possible “Hawthorne Effect”<sup>1</sup> by giving equal attention to all teachers and subjects.

Teacher participants assigned to the AVER or Kohlberg conditions worked with project personnel in up to eight day-long training sessions. Teachers assigned to each condition undertook differentiated training specific to the condition to which they had been assigned.

### *AVER Condition Training Sessions*

Five teachers met with project personnel for training sessions. Programmed text materials on evaluative reasoning (Meux, 1974) were utilized in the early sessions to elucidate the nature of value judgments, types of value judgments, moral and non-moral points of view, and the use of rudimentary procedure in the assembly of facts relevant to value judgments.

Teachers viewed a videotape of an AVER teacher demonstrating the use of syllogistic reasoning in classroom discourse. Teachers undertook similar tasks to those they would later administer to their students. Some emphasis was placed on the implementation of role-taking tasks in the classroom. Trainees were also instructed in the utilization of four methods of testing principles implied by moral judgments, the role exchange test, the universal consequences test, the new cases test, and the subsumption test.

Later training sessions involved examination and discussion of curriculum materials to be utilized in implementation of the AVER condition.

### *Kohlberg Condition Training Sessions*

Six teachers took part in Kohlberg training sessions. These sessions were of roughly the same duration and extended over the same months prior to the implementation period as the AVER condition training sessions.

Numerous tasks were undertaken by trainees under the direction of project personnel to establish a working knowledge of Kohlberg’s stages of moral

reasoning. Kohlbergian dilemmas were discussed with examples given of typical responses at each stage. Teachers were encouraged to compose their own dilemmas and to present appropriate arguments for choices at each of the moral reasoning stages.

Once the teachers were familiar with the basic stages of Kohlberg's theory of moral reasoning development, they were instructed in various methods of presenting dilemmas to groups of students. This instruction included the viewing of videotaped demonstrations in Kohlbergian moral education conducted by leaders experienced in dilemma discussion techniques. Later training sessions were devoted to the presentation of draft units of the Year 2 Kohlberg condition. Teachers and project personnel discussed potential implementation problems and suggestions for improvement of the curriculum materials.

*Evaluation Design and Subjects*

Table 1 shows the approximate number of students by treatment condition and school for whom complete pretest-posttest data were available on one or more of the measures. Differences in cell sizes within the schools and terms of the school year are related to the need to assign subjects to treatment groups of approximate "class" size within timetable blocks. Although it was possible to assign subjects randomly to treatment conditions within timetable blocks, it was not possible to assign them randomly across timetable blocks. The availability of project teachers within timetable blocks necessitated the unequal distribution of subjects to treatment groups shown in Table 1. The junior secondary design is incomplete in the sense that no non-values, control groups were available for comparison with the experimental conditions. Nevertheless, it will be noted that for all hypotheses the one experimental group served as control for the other. In the senior secondary school design, the possibility of comparing experimental groups with non-experimental control groups was also available.

TABLE 1  
NUMBER OF SUBJECTS BY TREATMENT CONDITION, GRADE, AND SCHOOL

School/Grade	AVER	Kohlberg	Control
Junior Secondary School: 3rd Quarter (Grade 8)	52	27	--
Junior Secondary School: 4th Quarter (Grade 8)	27	52	--
Senior Secondary School: 2nd Semester (Grades 11/12)	66	88	69

Table 2 shows differences in testing, treatment period, and intensity of treatment by school and school term. These differences persuaded the researchers to analyse separately data gathered in different terms and schools. In effect, then, this report describes findings of three discrete studies, junior secondary third quarter, junior secondary fourth quarter, and senior secondary, relevant to evaluation of the effects of the experimental conditions. It will be noted that the



TABLE 2  
TREATMENT PERIOD, TESTING AND INTENSITY OF TREATMENT BY SCHOOL

School/Term	Pretest	Treatment Period	Posttest	Intensity
Junior Secondary 3rd Quarter	February Week 1	February–March Inclusive	April Week 2	Concentrated (1 hour per day)
Junior Secondary 4th Quarter	April Week 1	April–May Inclusive	June Week 2	Concentrated (1 hour per day)
Senior Secondary 2nd Semester	February Week 1	February–May Inclusive	June Week 1	Distributed (50% experimental; 50% regular curriculum; 6 hours per week)

overall design does not permit the partialling out of effects attributable to concentrated versus distributed implementation of the treatment conditions.

Batteries of tests were administered during pretest and posttest periods. Monitoring of treatment implementation was done by means of direct observation by project personnel utilizing an observation checklist (AVER, 1975).

Test Development

A major task of Year 2 was the locating, modifying, and further development of measures of the various dispositions and abilities associated with moral competence.

In Year 2 of the project, the following measures were utilized in the pre and/or posttest batteries of the AVER-Surrey Project: the Defining Issues Test (Rest, 1974; modified by AVER, 1975), the Facts/Values Test (AVER, 1974; modified by AVER, 1975), the Normative Reasoning Test (AVER, 1974; modified by AVER, 1975), a Prejudice Test (Adams, 1975, Forms A and B), and the Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965).

The Defining Issues Test (DIT) is purported to measure moral development. Responses are keyed to Kohlberg’s stage types. The modified version of the DIT utilized by the AVER-Surrey Project differentiates between conventional and principled moral reasoning on five dilemma stories, including the Heinz, Escaped Prisoner, Newspaper, Webster, and Doctor’s Dilemmas. Rest claims that “P”<sup>2</sup>, or principled morality, scores based on the first three of these dilemmas correlate .93 with P scores based on the six dilemma long-form of the DIT. Evidence concerning the reliability and validity of the DIT is provided in Rest’s *Manual for the Defining Issues Test* (1974). It is worth noting that AVER modified the DIT by adding two dilemmas accompanied by response items judged to be more sensitive to lower stage (i.e., pre-conventional) reasoning than the Rest dilemmas and response items. However, P scores reported here are based only on the dilemmas and items developed by Rest.

The Normative Reasoning Test (AVER, 1975) utilized in Year 2 of the project comprised 30 items in which relationships between premises and conclusions in normative arguments were sound or unsound. Subtest A comprised 20 multiple-choice items for which pupils were required to select the soundest conclusions following from a value premise and a factual premise. Hoyt (1941)



estimates of reliability for samples of junior and senior secondary subjects in Surrey, B.C. schools range from .71 to .78 for subtest A. Subtest B comprised 10 multiple-choice items for which pupils were required to select the reason or premise that best completed a normative argument. Hoyt estimates of reliability for Subtest B range from .61 to .81. Overall estimates of reliability for the total 30-item test range from .78 to .86, according to Hoyt estimates of reliability, and .68 to .76, according to Cronbach's alpha (1951). For purposes of the group comparisons made in this study, such reliability coefficients were judged acceptable.

The Facts and Values Test (AVER, 1975) is a modification of the test used in Year 1. In pretest form, it comprised 20 items, but in posttest form it comprised 30 items that subjects were required to identify as either factual or value claims. Hoyt estimates of reliability range from .69 to .79 for four samples of junior and senior secondary school subjects in Surrey schools.

The Prejudice Test (Adams, 1975) utilized in the posttest battery is a 27-item Likert-type measure of subject's prejudicial attitude towards an ethnic group to whom considerable community hostility was being directed at the time of the study. Hoyt estimates of internal consistency ranging from .94 to .95 were obtained from samples of Surrey, B.C. junior and senior secondary school subjects on this measure.

The Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965) is a 26-item measure of a subject's disposition to deceive others about oneself. As such, it is a measure of frankness or honesty in self-reports. A high score indicates a high disposition to deceive. Hoyt estimates of reliability for the measure range from .76 to .80 for samples of Surrey, B.C. junior and senior secondary school subjects.

### *Year 2: Hypotheses and Findings*

The experimental/evaluation portion of Year 2 of the AVER-Surrey Project was designed to test certain hypotheses and to check for unintended effects of the experimental conditions on prejudice and disposition to deceive. Kohlbergian theory and Year 1 findings suggested the following hypotheses with respect to treatment group differences:

1. On the Normative Reasoning Test, it was predicted that  
AVER > Kohlberg = Control.
2. On the Facts and Values Test, it was predicted that  
AVER > Kohlberg = Control.
3. On the Defining Issues Test, it was predicted that  
Kohlberg > AVER = Control.

Recent literature on curriculum evaluation (Scriven, 1967; Worthen & Saunders, 1973) suggests that data should be gathered concerning unintended and undesirable effects of programmes. In the case of the AVER and Kohlberg conditions, possible negative effects of classroom discourse about prejudice were of concern. Hartshorne and May (1929-30) and others have reported evidence suggesting that some character training/moral education programmes increase rather than reduce pupil tendencies to deceive teachers and evaluators about their sentiments and conduct. In order to assess unintended treatment effects on these variables, the Prejudice Test (Adams, 1975) and Social Desirability Scale (Crandall, Crandall & Katkovsky, 1975) were administered to subjects. Results of

analyses of variance of treatment effects among senior secondary school subjects (the only portion of the design permitting experimental-control group comparisons) are reported here.

*Treatment Group Differences on the Normative Reasoning Test*

On the Normative Reasoning Test, it was predicted that AVER > Kohlberg = Control. Table 3 shows treatment group differences and an analysis of covariance of treatment effects on senior secondary school students. The finding of a significant treatment effect and of significant differences between the AVER and Kohlberg subjects and the AVER and control subjects, as well as the direction of those differences, tends to confirm hypothesis 1 for senior secondary students.

TABLE 3  
ANALYSIS OF COVARIANCE AND TREATMENT GROUP MEANS AND STANDARD DEVIATIONS ON NRT POSTTEST WITH NRT PRETEST SCORES USED AS COVARIATES: SENIOR SECONDARY SCHOOL SUBJECTS

Source	D.F.	M.S.	F	p
Treatment	2	214.28	9.35	.0001*
Error	199	22.92		
	Unadjusted $\bar{X}$	Adjusted $\bar{X}$	S.D.	N
AVER	23.95	23.13	] 3.66	61
Kohlberg	19.80	19.85		88
Control	19.24	20.08		54

] Significant ( $\alpha = .05$ ) treatment group differences according to Newman-Keuls Multiple Range Test.

\* Significant treatment effect.

Analyses of treatment effects of the Normative Reasoning posttest for junior secondary students in the 3rd and 4th quarters of the school year were confounded by significant treatment group differences on the pretest, despite random assignments of subjects to the AVER and Kohlberg conditions. When differences on the pretest were controlled by means of analyses of covariance, results revealed no significant treatment group differences among junior secondary school comparison groups.

*Treatment Group Differences on the Facts and Values Test*

On the Facts and Values Test it was predicted that AVER > Kohlberg = Control. However, analyses of covariance of posttest scores with pretest scores used as covariates revealed findings contrary to this prediction in one of the three populations sampled.

Table 4 shows data relevant to hypothesis 2 from junior secondary students in the fourth quarter of the school year. The difference, favouring as it does the Kohlberg condition, is contrary to hypothesis 2.



TABLE 4  
ANALYSIS OF COVARIANCE AND TREATMENT GROUP MEANS AND STANDARD  
DEVIATIONS ON THE FVT POSTTEST WITH THE FVT PRETEST SCORES  
USED AS COVARIATES: JUNIOR SECONDARY SUBJECTS, 4th QUARTER

Source	D.F.	M.S.	<i>F</i>	<i>p</i>
Treatment	1	153.89	7.49	.008*
Error	77	20.55		

	Unadjusted $\bar{X}$	Adjusted $\bar{X}$	S.D.	<i>N</i>
AVER	9.58	9.77	3.67	26
Kohlberg	12.83	12.74	5.13	54

\* Significant treatment effect.

*Treatment Group Differences on the DIT*

Analyses of covariance for the senior secondary and two junior secondary samples failed to confirm the prediction of significant treatment group differences in favour of the Kohlberg condition. In fact, with all three samples of subjects, the direction of mean differences favoured the AVER over the Kohlberg condition, although these may very well have been chance differences.

*Treatment Group Differences on the Prejudice Test (Adams, 1975)*

The prejudice unit comprised only a small portion of the experimental conditions in Year 2 of the AVER-Surrey Project. It is not unreasonable, however, to assume that moral education programmes dealing with the issue of prejudice might have some effect, either positive or negative, on prejudicial attitudes towards minority groups. In the interests of evaluating possible experimental-control group differences on this variable, a prejudice test was administered as part of the posttest battery to senior secondary school students.

High scores on the Prejudice Test indicate greater prejudice than do low scores. Table 5 shows a significant treatment effect, with the direction of mean differences suggesting that students in the control and Kohlberg conditions were more prejudiced than the AVER students. However, the Newman-Keuls test failed to confirm the significance of particular comparisons among the treatment conditions.

*Treatment Group Differences on the Social Desirability Scale*

The Social Desirability Scale (Crandall, Crandall & Katkovsky, 1965) can be interpreted as a measure of subjects' disposition to deceive others about themselves. As such, it taps a construct of great potential interest to researchers in moral education. Analysis of variance of treatment group differences in social desirability scores reveals no significant differences among comparison groups. This finding increases confidence that student self-reports on the Prejudice and Normative Reasoning Tests were not differentially affected by disposition to deceive.

TABLE 5  
ANALYSIS OF VARIANCE AND TREATMENT GROUP MEANS AND  
STANDARD DEVIATIONS ON THE PREJUDICE TEST:  
SENIOR SECONDARY STUDENTS

Source	D.F.	M.S.	F	p
Treatment	2	2686.83	3.97	.02*
Error	235	676.22		

	$\bar{X}$	S.D.	N
AVER	73.54	24.59	71
Kohlberg	81.41	25.09	91
Control	85.45	28.27	76

\* Significant treatment effect.

### *Monitoring of Conditions During Treatment Period*

A total of 54 observations of class sessions were made utilizing an observation checklist designed by project personnel (AVER, 1975). No estimate of interobserver reliability in the use of the instrument is presently available.

Although no tests of significance were applied to observation checklist data, some differences reported by project observers were consistent with those expected between AVER, Kohlberg, and control conditions. Descriptive analyses show that Kohlberg teachers posed dilemmas, encouraged reason-giving at a variety of stages, rephrased responses to clarify stage level, and refused to let pupils "escape the dilemma" 50% more frequently as a proportion of their teaching behaviours than did AVER teachers. The proportion of teaching behaviours in the facts/values distinction, syllogistic reasoning, and principle testing categories was 50% higher for AVER than for Kohlberg teachers. In the use of other categories of behaviour specified by the observation checklist, experimental and control teachers were roughly similar.

### *Conclusions and Implications*

A number of gains in research and practice in moral education have resulted or are likely to result from the findings and activities of the AVER-Surrey Project.

Moral education curriculum materials focusing on such issues as population, pollution, the aged, capital punishment and prejudice have been developed and evaluated. These materials are now being revised and edited in preparation for publication.

Teacher training sessions and classroom observations of teachers working with pupils provide, at least, pilot data concerning the effectiveness of our in-service training methods. Year 1 and 2 observations suggest that some teacher training methods were effective, whereas others were not. Differences among teachers trained in the same manner may be related to main effects and/or interactions of teacher characteristics, experimental condition, pupil characteristics, and teacher training methods. We cannot discount the possibility that not all teachers may be



suited to the teaching of certain types of programmes in values and moral education. Research into the effectiveness of teacher training methods in moral education and into differences among teachers in the various moral components is indicated.

The needs of research and theory in moral education have been advanced by the direct comparison of differentiated approaches to values and moral education *and* by the development, modification, and identification of tests suitable to assessing some of the effects of programmes in moral values education. Over the two years of the project, normative reasoning and facts-values tests possessing a moderately high degree of internal consistency have been developed. Tests of dispositions relevant to particular moral education programmes (i.e., the Prejudice Test and the Social Desirability Scale) have been identified, modified and utilized. The project has provided new data concerning the internal consistency of these measures. Although attempts to develop a valid and reliable measure of role-taking and to make the DIT sensitive to moral reasoning development from the preconventional to the conventional level have not yet reached fruition, considerable data have been gathered that may yet lead to advances in these areas of measurement.

The expected superiority of the AVER condition in promoting the ability to recognize sound and unsound normative arguments was partially confirmed. Among senior secondary school students, the AVER treatment was significantly more effective than the Kohlberg and control treatments in promoting this ability. It may be that the failure of the AVER condition to yield a significant treatment effect on the Normative Reasoning Test among junior secondary school students is related to developmental constraints on the acquisition of the skill being tested. However, it is also possible that chance or tester-induced systematic differences existing prior to the treatment period confounded the discovery of significant treatment effects.

Although not predicted, the finding of a significant AVER effect on the Prejudice Test is encouraging. It is possible that the AVER condition was more effective than the Kohlberg condition in promoting dissonance concerning subjects' prejudicial beliefs. However, such speculation must await further investigation, including replication of present findings.

Perhaps the most inexplicable finding of Year 2 is that of a significant treatment effect among fourth-quarter junior secondary students on the Facts and Values Test *favouring the Kohlberg over the AVER condition*. In Year 1, the Meux condition proved to have a significant effect on promotion of the facts-values distinction. On the strength of this finding, elements of the Meux programmed text materials were incorporated into the AVER condition for Year 2. Certainly the AVER treatment was the only Year 2 condition explicitly designed to enhance the skill of differentiating factual and value claims. However, it appears that the AVER condition might actually have depressed scores on the modified Facts and Values Test in Year 2 (!). A possible explanation is that the AVER program made subjects sufficiently sensitive to different classes of "factual" and value claims that the task presented in the test became more, rather than less, difficult to AVER subjects. However, this is merely conjecture. Certainly, the discrepant Year 1 and 2 findings concerning treatment effects on the Facts and Values Test warrant further investigation.

The expected superiority of the Kohlberg condition in promoting moral

reasoning development (as measured by the DIT) was *not* confirmed. The failure to yield significant results may be attributable to such factors as lack of power in the treatment, the relative insensitivity of the DIT to stage transition from the premoral to conventional level of moral reasoning or, even, extreme differences of cell sizes among the treatment groups analyzed.

The general lack of effectiveness of the Kohlberg condition and of the AVER condition among junior secondary school students requires further investigation. Did the insensitivity of the DIT as a measure of moral reasoning development at lower stages of Kohlberg bias the Year 2 results against the Kohlberg condition and/or was the form of the Kohlberg condition utilized in this study lacking in power? The ineffectiveness of the AVER condition on junior secondary students may well be related to the level of sophistication and/or difficulty of AVER materials and activities. The need to design AVER activities and materials specifically adapted to the reading and maturity level of younger students is certainly indicated by Year 1 and Year 2 findings.

This paper reports work conducted by the Association for Values Education and Research (AVER, Faculty of Education, the University of British Columbia) in a project funded by the Vancouver Foundation, the Educational Research Institute of B.C., the Surrey School District, and the Office of Field Development of the University of British Columbia Faculty of Education. The author presents this report on behalf of his colleagues and co-workers in AVER.

#### *Notes*

1. "Hawthorne Effect" may be defined as improved performance associated with participation of Subjects in situations involving the special attention accorded to experimental treatment groups (Sax, 1968, 339-340).
2. Rest claims that the P score, an index derived from responses to DIT items, can be interpreted as "the relative importance attributed to principled considerations" (Rest, 1974, 2-2).

#### *References*

- Adams, D. K. Prejudice test. Unpublished Likert-type attitude scale. University of British Columbia, Faculty of Education, 1975.
- AVER. *Report No. 5: Interim report of an experiment in moral education at Surrey, B.C.* Vancouver: Association for Values Education and Research, 1974.
- AVER. *Report No. 6: Final report of a study in moral education at Surrey, B.C.* Vancouver: Association for Values Education and Research, 1975.
- Coombs, J. R., & Meux, M. Teaching strategies for value analysis. In L. E. Metcalf (Ed.) *Values education: Rationale, strategies and procedures*. Washington, D.C.: National Council for the Social Studies, 1971.
- Crandall, B. C., Crandall, V. J., & Katkovsky, W. A children's social desirability questionnaire. *Journal of Consulting Psychology*, 1965, 29, 27-36.
- Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, 16, 297-334.
- Guidance Associates. *First things: Values*. Pleasantville, N.J.: Guidance Associates, 1972. (6 sound filmstrips and discussion guides.)
- Hartshorne, H., & May, M. A. *Studies in deceit* (Vol. I). New York: Macmillan, 1929-30.
- Hoyt, C. Test reliability estimated by analysis of variance. *Psychometrika*, 1941, 6, 153-160.
- Kohlberg, L. Stages of moral development as a basis for moral education. In C. M. Beck, B. S. Crittenden & E. M. Sullivan (Eds.) *Moral education: Interdisciplinary approaches*. New York: Newman Press, 1971.
- Metcalf, L. E. (Ed.) *Values education: Rationale, strategies and procedures*. Washington, D.C.: National Council for the Social Studies, 1971.
- Meux, M. & Associates. *Rational value decisions and value conflict resolution: A handbook for teachers*. Salt Lake City: Granite School District, 1974.

- Rest, J. *Manual for the defining issues test*. Minneapolis: University of Minnesota, 1974.
- Sax, G. *Empirical foundations of educational research*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.
- Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.) *Perspectives of curriculum evaluation*. A.E.R.A. Monograph 1. Chicago: Rand McNally, 1967.
- Selman, R. L. Level of social perspective taking and the development of empathy in children: Speculation from a social-cognitive viewpoint. *Journal of Moral Education*, 1975, 5 (1), 35-43.
- Turiel, E. Stage transition in moral development. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Williams, D. AVER in Surrey: An approach to research and development in moral education. *C.S.S.E. Yearbook*, 1975, 2, 62-75.
- Wilson, J. *Practical methods of moral education*. London: Heinemann, 1972.
- Worthen, B. R., & Sanders, J. R. *Educational evaluation: Theory and practice*. Worthington, Ohio: Charles A. Jones Publishing, 1973.



EVA SHAW

*The University of Calgary*

## The Effects of Attentional Focus on Graphic Discrimination

*This study compared the effectiveness of four methods of training young children in the skills of visual discrimination necessary to the acquisition of letter knowledge. Subjects were all children between the ages of four years and four years six months, attending day care centres subsidized by the City of Calgary, Alberta. Data were analyzed using a multivariate analysis of variance, with a probability of .05 as the level of significance. Within the limitations of the study, results indicated that training in oral description of critical cues of letters contributed to superior achievement in learning to match letters. In light of the findings of the study, an amendment to Gough's model of reading (Gough, 1972) was suggested. (Dr. Shaw is in the Department of Educational Curriculum and Instruction, Faculty of Education at the University of Calgary.)*

The transition from the understanding of language through audition to the understanding of language through vision is a complex process in which the eye must gradually supersede the ear as the organ of reception. In order to learn to read, it is essential that the learner be able to discriminate letters. Internalization of letters follows the hierarchical stages of matching (awareness of the sameness of two letters), identification (the ability to single out a specified letter from a group of letters on request), and recognition (the ability to write and name a letter on request from memory). In the process of learning to match letter forms, young children are asked to compare symbols which probably bear no associational meaning, and certainly bear no common associational meaning. Thus, the task is strictly perceptual. The child needs to be aware of line, slope, curve, and direction. He needs also to understand clearly the purpose of the task required of him. He cannot match unless he has a true perception of "same" and "different." As Gibson, Gibson, Pick and Osser (1962) pointed out, at no previous time in a child's life has the directionality of an object affected its label.

Recent research has shown an increased interest in the desire to describe clearly the process of interactions through which the reader absorbs information from print. Many of the subprocesses of reading appear to take place simultaneously and with greater or lesser effect on the end-product, according to the proficiency of the

reader. Studies of the complexity of the reading process have resulted in the construction of various models of reading. While there are differences among the models, most of them use single word recognition as the starting point of the reading process. Mackworth (1972), however, stated that if the letter learning itself has been inadequate, there is a strong possibility that the learning of simple word pattern and the linkage of these patterns with their sounds will present serious difficulties to the beginning reader.

The questions then arise as to how children begin the task of letter recognition. What, if any, scanning devices do they employ when making visual comparisons? What particular information do they select and retain from a symbol that may be totally devoid of meaning for them?

It is possible that some order may be brought into the child's perceptual environment through the use of language. There is general agreement that at some time in a child's development, the power of thought exceeds the power of language. Since language is the verbal expression of thought, it is obvious that what a child says, he is able to think. What is questionable is whether a young child's verbal capacity can be increased to express the complexity of his thoughts. Language was defined by Tansley (1970) as a "regulator of behaviour" and thus he saw perception, sharpened by language, leading to more effective conceptualization. The use of language leading to the verbal expression of strategies of comparison could provide the child with a means of reducing doubts in a manner which might reinforce his discriminatory skills.

### *Purpose*

Research evidence supports the importance of visual discrimination training in beginning reading. The main purpose of this study was to investigate the processes by which a young child discriminates visually, since this is the first step towards letter recognition, and to investigate the effect of a verbal response on visual discrimination. An attempt was made to gain knowledge of the invariant features by which prereaders distinguish letter patterns.

### *Background to the Problem*

Among recent models of the reading process the one by Gough (1972) presents a point of view emphasizing the need to investigate the process of letter recognition in the early prereading stages (Fig. 1).

The starting point of Gough's model is the eye fixation. Each fixation lasts approximately 250 milliseconds before the next saccade is made. As each fixation is achieved, the visual pattern reflected on the retina activates the visual system and results in the formation of an icon. Gough saw the icon as "an 'unidentified' or 'precategory' visual image" (Gough, 1972, p. 332) capable of including everything in an oval shape roughly two inches wide and one inch high. Reputable research findings concerning the existence and the function of the icon were accepted by Gough, but he drew attention to the minimum of knowledge on the formation of the icon. Available data led Gough to assume that the initial fixation results in an icon which, for practical purposes, would include 15 to 20 letters and spaces of a sentence. Recognition of the icon would be complete in about 100 milliseconds and decay of the icon would occur in about 250 milliseconds as it becomes replaced by the new icon resulting from the next fixation. As this activity is taking place,

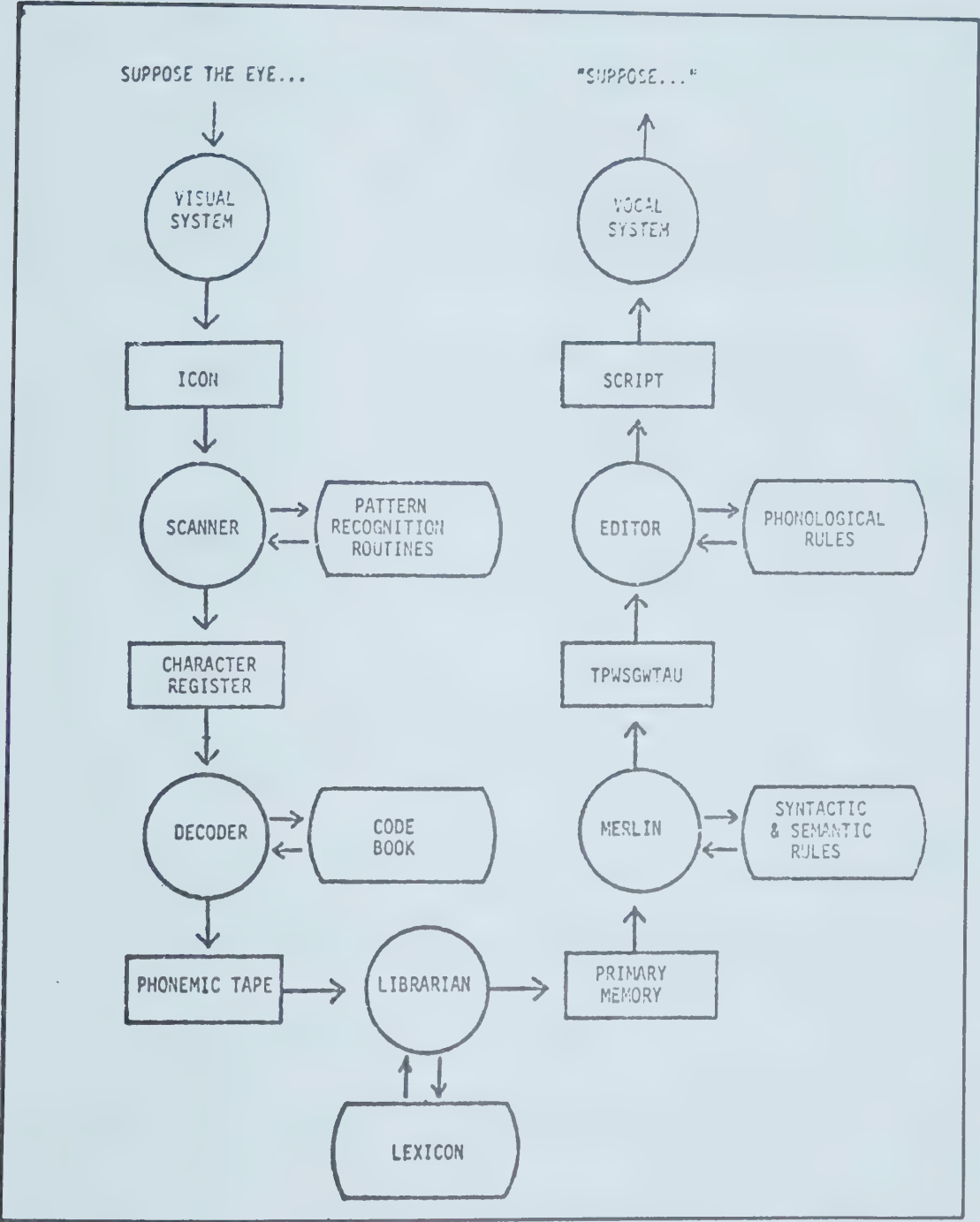


Figure 1. A model of reading. (From Philip B. Gough, "One Second of Reading" in J. F. Kavanagh and I. G. Mattingley [Eds.] *Language by Ear and by Eye: The Relationships Between Speech and Reading*. Cambridge, Mass.: M.I.T. Press, 1972. Reprinted by kind permission of the author and publishers.)

familiarity of the reader with the precategorical image results in identification. Gough assumed that the shapes are identified as letters.

When referring specifically to the prereader, Gough described him as being already equipped with several of the components needed for reading: a visual system which produces an icon, a vocabulary, a degree of comprehension, and a phonological system. However, what the child lacks is a scanner, a character



recognition device, and a decoder. Gough saw the decoder as a mechanism by which meaning becomes attached to the character.

Gough discussed the problem of character recognition and asserted that it was a problem, for while it had been shown that young children can discriminate letters, it was also obvious that children confuse letter shapes in successive presentations as they normally occur in the reading situation. There is no real information at this time about the features by which young children differentiate letters. Gough stated that if the subcomponents of the character recognition device could be identified, it might be possible to find out which subcomponents are present in the beginning reader.

The character register (Gough, 1972, p. 335) to which Gough referred appears to be a system of storing significant signals received serially from the icon, every ten or twenty milliseconds, in the form of letters. Signals, however, may emerge as letters only if the reader is able to label the signals as such. The fact that the prereader may have had no previous perceptual experience with the symbol confronting him makes it difficult for him to use retrieval mechanisms. Gough saw these mechanisms as the connecting link between the character register and the representations from print that are comparable with the reader's speech comprehension.

Since we do not know in what form the printed symbol, *per se*, appears in the icon, it is important to know how the critical features of letters are identified by a child. Gough appeared reluctant to categorize reversal and rotation errors as symptoms of reading disability; he accepted that they are not exclusive to poor readers. This suggests that accumulated experience in matching might characterize the method by which children learn to match. Such a position is supported by the statement that there is no evidence available to show that the teaching of discrimination is required since "the distinctive features of phonemes are not taught but they are nevertheless learned" (Gibson et al., 1962, p. 905). It appears, then, that knowledge of the relative strength of distinguishing features could affect the manner of teaching alphabet letters.

If, as Mackworth (1972, p. 716) stated, a word is maintained in the iconic store by the matching process, then presumably single letters would be similarly maintained, and the memory trace of the letter would be strengthened by the volume of matching until learning criterion of mastery is reached. When this criterion is reached may depend in part upon the child's memory storage system. As environmental information reaches the memory system, it enters primary or short-term memory. While the capacity of primary memory is limited, the process by which items pass from short- to long-term memory is considered to be under the possible control of the individual (Sticht, Beck, Hanke, James & Kleiman, 1974, p. 49). Information from the short-term memory may be retrieved accurately if it is retrieved rapidly. Speed of retrieval will depend upon the amount of preceding rehearsal and the strategies by which individuals relate new information to that already present in secondary or long-term memory. Sticht et al. (1974, p. 49) suggested that this information might be efficiently related by the linguistic recording of visually received information.

Since the capacity of primary memory is limited to approximately seven items, one item must recede as a new item enters primary memory. There may be certain conditions under which item recession is most likely to occur. Visual interference

may be one of these conditions; Barr (1972) suggested that a prereader can visually understand two to three letters at a glance, increasing by one letter per year. Since children's rates of development vary so widely, it may be that peripheral vision of nontarget letters interferes with a child's matching skill. Thus, the area of focus may control the reception of detail, and in visual discrimination it is the recognition of detail rather than the outline that ensures consistent accuracy. Geyer and Kolars (1973, p. 193) made reference to "an attentional focusing mechanism which delimits the size of the field", and it may be that training in the arousal of this mechanism could engender selective fixation, thus enabling a child to efface extraneous information.

There appears to be a need to investigate the possibility of a common surveying procedure in young children. It has long been known, for instance, that the upper half of lower-case letters is more easily identified than the lower half (Huey, 1908). More recently it has been established that stimuli in the left field of vision are more accurately perceived than those in the right field (Harcum, Hartman, & Smith, 1963), yet Kolars (1969) has shown the right side of lower-case letters to be more informational than the left side. Since the perceptual primary effect of the left field of vision appears to be strong and may be clearly differentiated from the response primary effect (Harcum, 1964), it could be that particular attention to the right side of a letter is warranted in developing skills of visual discrimination.

### *Hypotheses*

In order to investigate the effects of different methods of training young children in the skill of discriminating letter forms, the following null hypotheses were tested.

1. Subjects who receive training in matching a stimulus with its equivalent within an array of letters show no significant difference in letter matching ability from subjects who are trained to match letters singly. (Group VS)
2. Subjects whose attention during training is drawn to a specific field of vision show no significant difference in letter matching ability from those whose attention is drawn equally to the total field of vision. (Group VS-D)
3. Subjects whose attention during training is drawn to critical cues (those particular attributes of line, curve, and directionality which differentiate letters of the alphabet) by the investigator's verbal descriptions show no significant difference in letter matching ability from subjects who receive no such direction. (Group VB-D)
4. Subjects who receive training in the oral description of critical cues show no significant difference in letter matching ability from subjects who receive no such training. (Group VB-D-R).

### *Design of the Study*

This study was designed in four parts: screening, familiarization, training, and testing.

### *Subjects*

The sample for this investigation included children between the ages of four years and four years six months attending day-care centres subsidized by the City of Calgary, Alberta. A total of 132 subjects were assigned in random rotating order to one of four groups. There was no control group as such, but the treatment given to Group 1 (V-S) closely approximates the usual method of teaching letter discrimination to young children.



### Materials

*Screening.* A test of letter knowledge was given to control for prior experience. A simple test to determine the subject's understanding of the terms "same" and "different," and a test to determine the subject's initial awareness of the term "line," "curve," and "direction" were also administered.

*Familiarization.* Materials for this phase were constructed in exactly the same manner as were the training materials described below except that letter-like shapes were substituted for the letters used in training. Examples may be seen in Fig. 2. The purpose of this phase was to dispel any anxiety the subject may have felt, and to familiarize him with the expectations of the investigation.



Figure 2. Samples of letter-like shapes used in the study.

*Training.* Materials for all groups consisted of vinyl lower-case letters mounted on 8½" x 11" sheets of white paper. Letters were approximately 2" in height with a ⅜" width stroke. Twelve letters of the alphabet were used as stimulus letters in the training sessions. The letters were selected in such a way that characteristic letter features of line, curve, and direction would appear proportionately in both training and testing segments of the study. The stimulus letters were:

w t t v c z z a d d f h m p p r

The response letters were:

w a t v c y z a b d f h m g p r

For Group 1 (VS) the stimulus letter, printed in black, was placed on the left side of the page, while the response letter was one of a group of letters on the right side of the page.

For Group 2 (VS-D) the stimulus and response letters were mounted singly on opposing pages. All the letters were divided vertically in half. The left side of each letter was printed in grey, and the right side was printed in black. (Fig. 3)



Figure 3. Samples of letters used for Group 2 (VS-D).

For Group 3 (VB-D) and Group 4 (VB-D-R) materials were exactly the same. The stimulus and response letters, printed in black, appeared singly on opposing pages.

*Testing Materials.* All groups completed common tasks for testing purposes. For the first test (Letter Discrimination I), the following 14 lower-case stimulus letters appeared, centred singly on 8½ x 11-inch sheets:

l o i x k s y b e g j n q u.

The following lower-case response letters appeared on a further 14 sheets:

l a i x f s v b e g j m q n.

The sheets were arranged in a binder so that the appropriate stimulus/response



pairs appeared opposite each other. For the second test (Letter Discrimination II), 14 sheets were prepared, each sheet displaying 4 lower-case letters. The first letter on each sheet was the stimulus letter, while the remaining three served as response choice letters. The same stimulus letters were used as in Letter Discrimination I. For the final test (Letter Preference), a set of 18 sheets was prepared. Three different lower-case letters were spaced on each sheet. The following groups of letters were used: “o c e,” “f t h,” “d a p,” and these groups were presented in all possible orders.

Procedures

Screening, familiarization, training, and testing treatments were given individually to each subject at one sitting. Each group received a different kind of training. In all training groups, sixteen responses constituted one learning trial and each subject received two consecutive learning trials.

TRAINING GROUP	TREATMENT
Group 1 (VS)	Visual attention, undirected
Group 2 (VS-D)	Visual attention, directed
Group 3 (VB-D)	Visual attention, verbally directed
Group 4 (VB-D-R)	Visual attention, verbally directed with oral response

Figure 4. Treatments for training groups.

Training Procedures

*Group 1 (VS).* Subjects were asked to look carefully at the stimulus letter on the left and point to the matching letter from the group on the right side of the page. Errors were corrected as they occurred by showing the subject the correct response.

*Group 2 (VS-D).* Subjects were asked to look carefully at two-tone letters displayed on the opposing sides of the binder and state whether they were the same as each other. Errors were corrected as they occurred.

*Group 3 (VB-D).* For each subject in the third group, the binder was opened displaying two black letters on opposing pages. The subject was asked to judge if they were the same as each other. If the subject responded correctly, the investigator said, “Good. Both letters have straight lines at the left, a curve going over the top, and another straight line at the left.” The investigator’s comments varied in specific content, but always confirmed correct responses in terms of critical cues. Error responses were corrected as they occurred by drawing the subject’s attention to differences or similarities using terms of critical cues.

*Group 4 (VB-D-R).* For each subject in the fourth group, the binder was opened displaying two black letters on opposing pages. The subject was asked to judge if they were the same as each other. If a correct response was given, the investigator asked, “Why are they the same/not the same?” The response was recorded on a sheet designed to categorize expected responses. Twelve categories were prepared, and each category as subdivided into two columns, V and E, indicating whether the response was volunteered or elicited. If the subject’s reply did not contain reference to at least two critical differences or similarities between letters, the investigator

then asked two or three of a set of previously prepared leading questions such as, "Do they both have straight lines at the left side?", "Do they both curve over to the right?" A pencil was used as a pointer to indicate directionality. When difference or identity had been established, the subject was then asked again to say why the letters were the same/not the same. These responses were reported as elicited. If the subject's reply did contain two or more references to differences or similarities, the response was reported as volunteered. Correct responses were verbally confirmed and, if a particular point of sameness or difference had been omitted, the investigator used suitable questions from the prepared set to complete the comparison. The investigator then said, "Now tell me another reason why they are the same/not the same." Resulting responses were considered as elicited responses.

### *Testing Procedures*

There was a common testing procedure for all subjects in all four groups immediately following the training period.

*Test 1 (Letter Discrimination I).* Subjects were required to compare 14 pairs of letters and state if they were the same or not the same as each other.

*Test 2 (Letter Discrimination II).* The subject was asked to look carefully at the letter on the left side of the sheet and point to its equivalent within the group on the right side. Fourteen responses were required.

*Test 3 (Letter Preference).* The subject was asked to look carefully at a group of three different letters and to select the two letters that looked most nearly the same. Three groups of letters were presented in all possible orders, thus calling for 18 responses.

### *Analysis of Data*

A multivariate analysis of variance was used for the four types of treatment. After the MANOVA, the differences between means were investigated by tests. A transformation of raw scores was performed to determine the possible effect of letter knowledge and awareness of critical cues on the dependent variables. Data obtained from Test 3 (Preference Test) were subjected to a chi square test.

### *Findings and Conclusions*

Tables 1 and 2 show values obtained for differences among all pairs of group means for correct responses in training trials and in Letter Discrimination I and II.

### *Main Effects*

The treatment in which the subject was required to make an oral response in terms of line, curve, and direction was shown to be significantly superior to each of the other three treatments in developing skills in letter-matching.

*Effects — Group 4 (VB-D-R).* Since it was assumed that all four groups were of approximately equal learning ability, it may be posited that some of the achievement differences could be due to the attentional energy control applied by the investigator during the training of this group. Volunteered verbal responses indicated whether or not the subject was looking in the appropriate place, whether he had a purpose for looking, and whether he was conserving his attentional energy for expenditure on relevant attributes. Thus, the investigator could direct questions in the most economical and productive manner. In recording the

TABLE 1  
T VALUES OBTAINED FOR DIFFERENCES AMONG ALL PAIRS OF GROUP MEANS  
FOR CORRECT RESPONSES IN TRAINING TRIALS 1 and 2

Test	Treatment Group	2 (VS-D)	3 (VB-D)	4 (VB-D-R)
1	1 (VS)	1.08	3.51**	3.39**
1	2 (VS-D)		.92	2.16*
1	3 (VB-D)			1.25
2	1 (VS)	2.09*	3.12*	5.31***
2	2 (VS-D)		.31	2.79**
2	3 (VB-D)			6.34***

\*  $p < .05$   
 \*\*  $p < .01$   
 \*\*\*  $p < .001$

TABLE 2  
T VALUES OBTAINED FOR DIFFERENCES AMONG ALL PAIRS OF GROUP MEANS  
FOR CORRECT RESPONSES IN TESTS 1 and 2

Test	Treatment Group	2 (VS-D)	3 (VB-D)	4 (VB-D-R)
1	1 (VS)	0.0	1.49	6.99*
1	2 (VS-D)		0.665	7.21*
1	3 (VB-D)			7.07*
2	1 (VS)	0.377	0.436	5.25*
2	2 (VS-D)		0.174	5.55*
2	3 (VB-D)			6.26*

\*  $p < .001$

responses of these subjects, it was observed that in the second training trial there were noticeable increases in 11 of the 14 categories, and that all of these increases were volunteered as opposed to elicited responses. One of the largest increases appeared in the category recording directionality responses. All subjects in the study, except for one in the fourth group, identified "b" and "d" as being the same in the first learning trial. However, the fourth group (VB-D-R), subjects who gave an oral response, made an improvement of approximately 60 per cent on the second trial, while the second group (VS-D), subjects whose attention was directed to a specific field of vision, gained approximately 27 per cent. Improvement shown by the remaining two groups in discriminating "b" and "d" was minimal. This finding supports the view of Jackson (1972) that "confusions resulting from mirror-image type stimuli may be implicit in the stimuli and can be viewed as normal rather than abnormal," yet at the same time supports the view of Hendrickson and Muehl (1962) who suggested that research attention be directed not solely towards what children do, but towards what children can do after effective training.

*Effects — Group 3 (VB-D).* The subjects in this group had their attention drawn to critical cues of letters by the investigator's verbal descriptions of



similarities and differences. This group performed at a significantly higher level than the first group (VS) in both training trials, but did not maintain this superiority in letter matching. The onset of habituation may provide a possible explanation for such deterioration in performance. The condition of habituation is a normal part of the learning process, for when a particular piece of learning is complete, habituation takes over and frees the attention for the processing of new information. However, since the verbal descriptions were repetitive to some extent, they may also have become monotonous stimuli which failed to sustain the required level of alertness. If this was the case, habituation may have occurred before a sufficiently high level of learning was attained, thus accounting for the reduced number of correct responses. The sole difference between treatments for this group (VB-D) and the fourth group (VB-D-R) was that overt verbal participation was required of the fourth group. Thus it may be that the verbal response is a motivational buffer that prevents orientation from becoming habituation before the optimal time.

*Effects — Group 2 (VS-D).* Subjects in this group had had their attention drawn to a specific field of vision by means of an increase in the intensity of colour at the right side of the stimulus. A group of z scores revealed the positive effect of practice on the second training trial, followed by a dramatic decline in performance on the first letter discrimination task, where there was a 12 per cent decrease in the number of correct responses. This decrease indicates that, in training, subjects were possibly being directed towards the right field of vision, but could not function later at the same level of efficiency without the focal guidance provided by the black contrast.

*Effects — Group 1 (VS).* Subjects in this group had been required to match a single stimulus with its equivalent within a set of letters; these subjects achieved the lowest number of correct responses in both training and testing. The findings indicate that the subjects may have had too much opportunity for error and too little positive reinforcement for correct response. It is suggested that undirected training has insufficient transfer value for maximum learning to occur. It is also possible that the subjects' perception may have been influenced by an excess of peripheral stimuli for them to have made the appropriate selection without guidance towards relevance and purpose.

#### *Letter Preference Test.*

Subjects were asked to select from a group of three letters the two which looked most nearly alike. The response choices were "o c e," "f t h," "d a p." Chi square tests applied to the data showed preferences to be significantly beyond the chance level. In this study there were 498 selections of "c o," 736 selections of "t, f," and 535 selections of "p d," indicating that the curve in a letter, rather than its line or directionality, is a most salient feature in young children's strategies of comparison. If all subjects in the study had made the same selection with total consistency, the number of selections would have been 792. Results thus show a marked consistency of choice, the consistency being highest in the fourth group (VB-D-R). The omission rate also was highest in the fourth group, where it appeared that, for some subjects, the mind set for identity was too strongly embedded to permit a switch to similarity.

*Implications of the Study*

In this study it was shown that the level of letter knowledge within the groups was not a contributing factor to ability in letter matching; thus it appears that abstract expressions were used to distinguish unnamed stimuli. On the other hand, it was observed frequently throughout the study that many subjects without letter knowledge already had names for the stimuli. For instance, "m" was referred to as "the bridge," and "r" as "the gun." It is possible that some of these labels may have developed as a result of exposure to alphabet books and blocks, some of which, in an effort to make early learning attractive, obscure the critical details in decorative embellishment. Since this type of labelling is too specific to have transfer value, it is suggested that attention be paid to correct labels.

On some occasions it was noted that the subjects were comparing the centres of the stimuli rather than the outlines. Since the width of the letter stroke was  $\frac{3}{8}$  inch, it appeared that this width was affecting the subject's ability to separate field from ground. Traditionally, a young child's formal introduction to the alphabet is through the use of large, thick letters. Assuming the letters are without meaning to the child, they must simply be seen as a pattern or series of markings. Hilgard (1957, p. 377) has pointed out that "patterns do not have to contain identifiable objects to be structured as figure and ground." This statement, combined with observations from the present study, would warrant questioning the practice of introducing letters only in a magnified form not usually found in reading books. With constant exposure to the magnified form, the child may be making a perception contrary to the one being reinforced by the teacher.

Observations led the investigator to believe that there exist a number of prerequisites to competency in letter matching. The child must know what a line is, and then must be able to compare lines in terms of height, slope, and relative position. He must have a spatial vocabulary in order to make meaningful reference to letter attributes and he must have a way of expressing his recognition of orientation.

The act of comparison appears to be refined in three stages: (1) recognition of gross difference, (2) recognition of fine difference, and (3) recognition of similarity. Thus, until children understand sameness, they cannot confidently express difference.

*A Re-examination of Gough's Model*

The model presented by Gough has demonstrated the process of the reading act in minute detail. The particular portion of the model towards which the attention of this study was directed illustrates the process by which the stimulation of the visual system forms an icon which then activates the scanner. The scanner then interrelates with pattern recognition routines before reaching the character register.

Gough introduced his article in this manner: "Suppose the eye of a moderately skilled adult reader . . ." (Gough, p. 331). Thus, he made clear his basic assumption of a degree of reading ability. However, in later discussions of the beginning reader, Gough attributed to the child all the factors necessary for "the assembly of . . . a primitive reading machine," lacking only a scanner. It is likely that all seeing people possess a scanner, and that in control of the scanner lies facility in visual discrimination. Children who have full letter knowledge have learned, within



varying degrees, to control the scanner. Children without letter knowledge must learn that control. The free scan of the human eye needs to be directed so that it functions as does an optical scanner in an electronic computer system, picking up only relevant information. The results of the present study indicate that scanner control may be facilitated through the use of oral language.

It may be noticed that Gough showed the processing sequence flowing from the scanner to the character register. Beginning readers are probably, although not necessarily, able to employ the character register as a place where letters or words appear as commonly named entities. This is not likely to happen before the child has reached the stage of letter recognition, and yet there is no reason to assume that the portion of Gough's model under discussion could not be applied to all activities of conscious looking. Therefore, it is suggested that an insertion be placed between the icon and the character register which would account for the processes involved in graphic discrimination. This proposed insertion is shown in Figure 5.

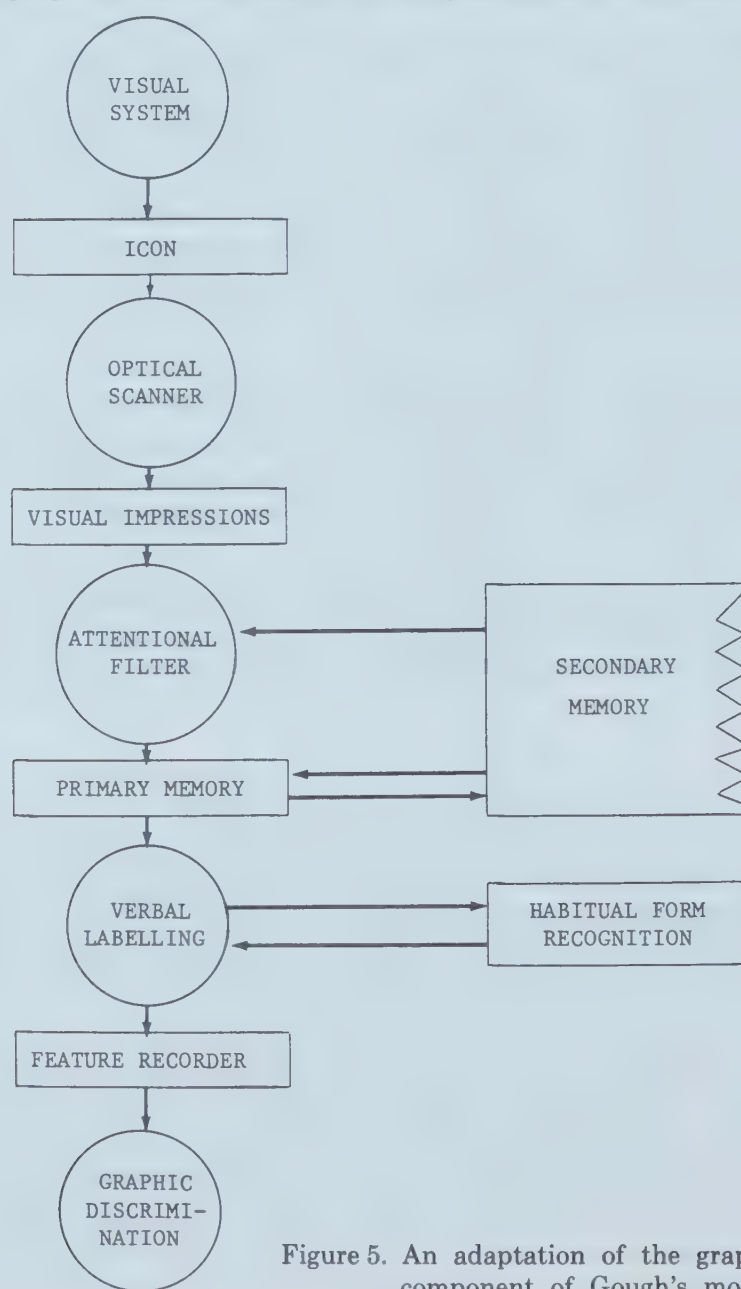


Figure 5. An adaptation of the graphic discrimination component of Gough's model of reading.



In this figure of a proposed model, the sequence moves from activation of the optical scanner to the reception of visual impressions. These impressions may be multitudinous and are of extremely brief duration. Because of their short duration the impressions must be sorted, categorized, or coded immediately in order to retain only meaningful information. Thus, a child needs to know that under given circumstances particular attributes have high relevance, and perhaps more particularly, some attributes have no relevance. In letter matching, for instance, size and colour may be irrelevant. In sorting visual impressions, a child uses his limited supply of attentional energy. Since there is a developmental aspect to the processing of attention, differences in the quality of attentional control may be expected from children of the same chronological ages. Sticht et al. (1974, p. 51) discussed developmental changes in terms of the gradually lessening degree to which the environment rules the organism. They referred to the alteration from "... the passive registration of information in displays to the active pursuing and processing of information" as the crux of the difference between seeing and looking.

Thus, much of the input of visual data is lost because it cannot be attended to before decay occurs. The visual impressions apparently pass through an attentional filter, leaving behind only those impressions which have been solidified in some manner. One of the most probable manners for solidification to occur may be through the act of labelling. By this act, features or groups of features may be clustered together under a single reference. The reference at this point need not be universally acknowledged, but must be meaningful to the individual.

The most vigorous impressions will then be taken into the primary memory. It may be recalled that Gough placed primary memory much further on in the sequence than it appears in the proposed model (Fig. 5). He defined primary memory as "a small capacity buffer storage system where four to five verbal items are maintained for a matter of seconds." (Gough, 1972, p. 339). Items within the primary memory may be ignored, or may be renewed through rehearsal. As each new item enters, it displaces an item already held within the primary memory. However, it is suggested in the proposed model that primary memory is involved in the sequence as soon as the organism becomes active beyond the initial point of visual stimulation. Gough's early placement of pattern recognition routines implies rehearsal, for routines cannot exist without rehearsal which must surely have an effect on primary memory. It is not suggested here that Gough's placement of primary memory is incorrect in the given position. The suggestion is that primary memory is active earlier in the sequence and continues to be active throughout the sequence.

Gough's (p. 340) reference to secondary memory was brief, and as a separate entity this factor did not receive prominence in his model. Smith (1971, p. 78), however, stressed that:

... nothing could be identified ... if a contribution were not made by long-term memory, because it is there that is lodged the knowledge of the world to which all incoming information must be referred.

On this basis the proposed insertion places secondary memory on a level with the attentional filter and primary memory to indicate that constant enriching interaction takes place between these factors, and this interaction may lead to habitual form recognition as a result of combining old knowledge with new. The

broken edge of the block in Fig. 5, representing secondary memory, suggests the lack of limit to the depth from which old knowledge may be drawn.

The next step in Gough's model is the character register, but nothing can be registered until it has a name. The name may be vague, perhaps as vague as "that thing" as opposed to "that other thing," for even this weak description implies a degree of discrimination. It is indicated in the proposed model that before graphic discrimination occurs, salient features of the stimuli are recorded, thus leading to discrimination. Only after discrimination would a character be registered as such. It is likely that the fineness of the discrimination will, for a young child, be related to the precision of the verbal label.

#### References

- Barr, R. Perceptual development in the reading process. In R. Hodges & H. Rudolf (Eds.), *Language and learning to read: What teachers should know about language*. Boston: Houghton-Mifflin, 1972, 131-139.
- Gibson, E. J., Gibson, J. J., Pick, A. D., & Osser, H. A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology*, 1962, 55 (6), 897-906.
- Geyer, J. J., & Kolers, P. A. Reading as information processing. In M. Voigt (Ed.), *Advances in librarianship*, 4. New York: Academic Press, 1973, 175-237.
- Gough, P. B. One second of reading. In J. F. Kavanagh & I. G. Mattingley (Eds.), *Language by ear and by eye: The relationship between speech and reading*. Cambridge, Mass.: Massachusetts Institute of Technology Press, 1972, 331-358.
- Harcum, E. R. Effects of symmetry on the perception of tachistoscopic patterns. *American Journal of Psychology*, 1964, 77 (4), 600-606.
- Harcum, E. R., Hartman, R. R., & Smith, N. F. Pre- versus post-knowledge of required reproduction sequence for tachistoscopic patterns. *Canadian Journal of Psychology*, 1963, 17 (3), 264-273.
- Hendrickson, L. N., & Muehl, S. The effects of visual discrimination pretraining with word and letter stimuli on learning to read a word list in kindergarten children. *Journal of Educational Psychology*, 1962, 52 (5), 236-241.
- Hilgard, E. R. *Introduction to psychology* (2nd ed.). New York: Harcourt, Brace, 1957.
- Huey, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908 (reprinted 1920).
- Jackson, M. S. Modes of adaptation to the first confrontation with English orthography in the visual modality. *Journal of Learning Disabilities*, 1972, 5(1), 29-34.
- Kolers, P. A. Clues to a letter's recognition: Implications for the design of characters. *Journal of Typographic Research*, 1969, 3 (2), 145-167.
- Mackworth, J. F. Some models of the reading process: Learners and skilled readers. *Reading Research Quarterly*, 1972, 7 (4), 701-733.
- Smith, F. H. *Understanding reading*. New York: Holt, Rinehart and Winston, 1971.
- Sticht, T. G., Beck, L. J., Hanke, R. N., James, J. H., & Kleiman, G. M. *Auditing and reading: A developmental model*. Alexandria, VA: Human Resources Research Organization, 1974.
- Tansley, A. E. Readiness for reading — Can it be hastened? In M. Chazan (Ed.), *Reading readiness*. Swansea, S. Wales: University College of Swansea, 1970, 70-74.



GLEN S. AIKENHEAD

*University of Saskatchewan*

## Using Qualitative Data in Formative Evaluation\*

*Formative evaluation requires sensitive techniques to assess student learning. This article explores a technique for: (a) acquiring qualitative data from standardized instruments, and (b) using these data in the formative evaluation of new curricula. Assets and limitations to the qualitative analysis technique are investigated. A new tenth grade course, Science: A Way of Knowing, was evaluated with respect to its impact on students' learning the processes, the nature, and the social aspects of science. The qualitative analysis technique proved to be a valuable tool to the developers of Science: A Way of Knowing because it yielded information that led directly to course revisions. (Dr. Aikenhead is Associate Professor in the Department of Curriculum Studies, College of Education, University of Saskatchewan.)*

The general purpose of this study is to illustrate a technique for evaluating a science curriculum project. The study explores the role which qualitative data can play in formative evaluation. Assets and limitations to the qualitative analysis technique are clarified in terms of its usefulness to curriculum developers.

Qualitative data are found by asking such questions as: What ideas have students learned? What misunderstandings have they acquired or retained? The essential advantage to using qualitative achievement data is in their power to reduce ambiguity in the evaluation process.

Test scores can be very ambiguous. The following claim is typical: "Students of course E gained 3.56 points on test T while students of course C gained only 1.71 points." These gains may be statistically significant, but they give the curriculum developer and teacher no indication as to what students actually learned. Test scores by themselves have very limited meaning (Aikenhead, 1973; Lucas, 1975).

Qualitative data are derived by analyzing test items and noting the content of those items. The items may be selected simply because of a high number or a low number of correct responses they received from students. (Examples are cited below.) For a pretest posttest design, one can identify items for which students

---

\*Based on a paper presented to the 51st annual meeting of the National Association for Research in Science Teaching, Toronto, Canada, April 2, 1978.



made a significant gain or loss during a period of instruction. Consequently, student gains in understanding and in misunderstandings are identified. This type of item analysis is accomplished with the McNemar (1969) chi squares statistic. Again, examples are cited below.

Only a few studies have taken advantage of the articulation of student learning generated by qualitative data. Cooley and Klopfer (1963) first demonstrated its effectiveness when they evaluated Klopfer's *History of Science Cases*. The McNemar chi square indicated test items which showed significant student gains. The content of those items comprised the qualitative data needed for the summative evaluation of the case studies. Klopfer and McCann's evaluation of a unified science course lent further credence to the usefulness of qualitative data. They identified weak areas in student understanding and made specific recommendations to teachers concerning what needed improving in the course (for example, "the aims and activities of science"). MacKay (1971) studied the common misunderstandings among seventh to tenth grade students in Australia. He was able to indicate which false notions seemed to be unchanged over those years. Similarly, Tamir (1972) documented some major misconceptions which Israeli high school students held about the nature and processes of science. It was hoped that these qualitative data would influence the science instruction in Israel. *Project Physics* was evaluated in terms of knowledge acquired about science and scientists (Aikenhead, 1974b). The McNemar chi square item analysis pointed to a specific knowledge for which students showed a significant increase or decrease in understanding. Thus the *Project Physics* and control groups were compared in a manner amenable to teacher interpretation. For example, *Project Physics* students generally gained in their understanding of the tactics of science, its values, its institutional functions, and the interaction of science with society, while the control group made almost no improvement in these areas at all. In the same arena of knowledge, Aikenhead (1975a) evaluated the Saskatchewan Chemistry program by: (1) listing specific misconceptions that were generally held by chemistry students, and (2) assessing the program's impact on students' knowledge about science and scientists. This assessment relied on qualitative data generated by the McNemar statistic. The findings of that study led to a major revision in the province's chemistry syllabus (Aikenhead, 1977).

In the summative evaluation projects described above, qualitative data played a supporting role. However, qualitative data could likely serve a more crucial function in the evaluation process when used in formative evaluation (Aikenhead, 1974b). This function has not been critically explored. The present study investigates the usefulness of qualitative data in formative evaluation.

### Procedure

#### Treatment

A new kind of science course, *Science: A Way of Knowing* (Aikenhead & Fleming, 1973), is presently being developed in Saskatchewan, Canada, for tenth grade students of average ability. The full year course is an activity-centered, interdisciplinary approach to learning *about* science and scientists: the aims of science, its epistemology, its tactics and strategies, its values, its interaction with society, and its human characteristics. *Science: A Way of Knowing* allows for a high degree of teacher flexibility in the content, sequence, and instructional methods (Aikenhead, 1975b).

### Field Trials

In 1974, four Saskatchewan teachers (three urban, one rural, total number of students = 141) volunteered to teach *Science: A Way of Knowing* draft #2. Feedback was gained through parent and student questionnaires, teacher and student interviews, and achievement tests. As a result, draft #2 was polished by making a number of small changes such as using appropriate cartoons, rewriting some thought provoking questions, and adding more optional activities.

In 1975, four teachers field tested draft #3 (two urban, one rural, and one — the author — in a Swiss International school). Only one teacher (urban) participated in both field trials. The total number of students was 180.

The study reported here is part of the formative evaluation of *Science: A Way of Knowing* draft #3. The developers wished to know:

- a. What ideas have students learned *about* science and scientists?
- b. What misunderstandings have they still retained?
- c. How do the answers to the above questions vary when the course is taught by different teachers?

These questions lend themselves to a one-group pretest posttest design using the McNemar chi square item analysis. At this stage in the development of *Science: A Way of Knowing*, there is no intention of comparing treatment groups with control groups. The teacher effect on student achievement is taken into consideration through the comparison of draft #3 results with draft #2 results.

### Instruments

To assess knowledge *about* science and scientists, students wrote the *Science Process Inventory*, SPI (Welch, 1969), and the *Test on the Social Aspects of Science*, TSAS (Korth, 1968). A detailed description of these instruments, including their validation and their use in research, may be found elsewhere (Aikenhead, 1973). The reader is encouraged to note some of the philosophical assumptions inherent in these instruments, as described by Lucas (1975).

The SPI deals with a student's awareness of the activities, assumptions, products, and ethics of science. It has 135 items. The form of the SPI usually used by researchers, form D, has a forced choice response format. Students must agree or disagree with each item. In an earlier study evaluating *Project-Physics*, Aikenhead (1972, p. 221) found that the forced choice format led to ambiguities for the researcher. One did not know if a student was disagreeing with a statement because: (a) she really did disagree, or (b) she did not understand what the item meant. Therefore the agree/disagree format of form D was replaced by a Likert-type scale: 1. strongly agree, 2. agree, 3. do not understand the meaning, 4. disagree, and 5. strongly disagree. The response "do not understand the meaning" allows students to indicate that a choice is impossible due to vocabulary or syntax. A student is forced to take a stand so long as she thinks she understands the meaning of the item. In order to distinguish between the two response formats, the Likert-type response instrument will be called "SPI form DL." Because students are not forced to guess with the Likert-type scale, the average scores should be lower for form DL.

The TSAS deals with: the interactions among science, technology and society; the social nature of the scientific enterprise; and the social and political responsibilities of scientists. There are 52 items. The original five-point Likert



scale was modified slightly. The mid response “uncertain or undecided” was changed to “do not understand the meaning.” Again, this was done to force students to take a stand if they felt they understood the item.

Both tests were marked in the same manner. Each item has a “correct” response based on the content validity of the instrument. The response “agree” or “strongly agree” or the response “disagree” or “strongly disagree” was considered correct if it concurred with the test’s key. The “do not understand the meaning” response was always marked incorrect, however it was not ignored. Each item was later analyzed with respect to the percentage of students who agreed with, did not understand, or disagreed with, the item.

*Data Analysis*

The usual test score information was generated in two ways: (1) for each of eight teacher groups, and (2) for each of two combined groups, draft #2 and draft #3. A *t*-test for matched pairs determined the significance of gains between the pretest and posttest scores, for both the SPI form DL and the TSAS.

Using the draft #3 combined group data, each test item was analyzed by the McNemar (1969) chi square statistic.

$$\chi^2 = \frac{(A - D)^2}{(A + D)}$$

with 1 degree of freedom.  
where A and D are cell frequencies  
in the following contingency table.

		Posttest	
		right	wrong
Pretest	wrong	A	B
	right	C	D

This statistic identified test items which experienced a statistically significant increase or decrease in student response between the pretest and posttest.

Each test item was also analyzed to determine the percentage of students making one of the three responses: agree, do not understand the meaning, and disagree. (The percentage of students not responding at all was also noted.)

*Results and Findings*

This report deals with the formative evaluation of *Science: A Way of Knowing*, draft #3. Occasionally, information from draft #2 is included for reasons of comparison only. Student responses to the SPI form DL and TSAS were analyzed both quantitatively and qualitatively.

*Quantitative Analysis*

Table 1 contains standard data for the SPI form DL and TSAS scores. The gains between the pretests and posttests are statistically significant. Students’ achievement was slightly greater for the SPI form DL than for the TSAS. (The actual gains are particularly large in comparison with others reported in the literature.) In spite of a change of teachers and in spite of an “improvement” in the instructional materials, the gross total test scores seem stable over drafts #2 and #3.



TABLE 1  
STUDENT TEST RESULTS FOR TWO DIFFERENT FIELD TRAILS,  
TEACHER GROUPS COMBINED

	Draft #2 — 1974-1975						Draft #3 — 1975-1976					
	SPI Form DL (N=137)			TSAS (N=109)			SPI Form DL (N=171)			TSAS (N=149)		
	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain	Pre- test	Post- test	Gain
Mean score <sup>a</sup>	85.0	96.0	11.0	27.4	31.1	3.7	85.4	96.5	11.1	28.0	31.7	3.7
Standard deviation	13.1	12.4		5.4	7.1		11.3	12.2		5.4	6.7	
Reliability <sup>b</sup>	.85	.85		.65	.78		.80	.84		.67	.76	
t test (matched pairs)	10.9	p < .001		5.4	p < .001		15.2	p < .001		8.92	p < .001	

<sup>a</sup> Total possible scores are: SPI, 135; TSAS, 52.

Typical mean scores for Saskatchewan 11th grade Chemistry students are: SPI Form DL, 92; TSAS, 28.

<sup>b</sup> K-R formula - 20 coefficient.

Information based on teacher groups is shown in Table 2. There is variation among the teacher groups, especially for the TSAS scores. However, other studies reveal even greater variation among teacher groups (Aikenhead 1972, 1975a). These results support the cliché that the individual teacher has more effect on student achievement than a particular curriculum. (The idea behind *Science: A Way of Knowing* is to give support to teachers having interests which correspond to the objectives of the course.) The predominant “no significant difference” for TSAS total scores suggests that students are not learning much about the social aspects of science. However, this is where quantitative data are most limited in giving useful information to a curriculum developer.

TABLE 2  
SUMMARY OF STUDENTS' TOTAL SCORES GROUPED BY TEACHER

Teacher	SPI Form DL						TSAS					
	N	Pretest	Posttest	Gain	t-test <sup>a</sup>	Proba- bility	N	Pretest	Posttest	Gain	t-test	Proba- bility
<u>Draft #2</u>												
A	76	86.1	98.4	12.3	10.4	p<.001	43	26.3	43.1	7.8	8.1	p<.001
B	21	77.2	91.0	13.8	5.6	p<.001	22	26.0	27.0	1.0	0.8	not sig.
C	16	88.6	98.5	9.9	2.7	p<.02	15	31.4	34.4	3.0	1.8	not sig.
D	24	85.7	91.0	5.3	1.7	not sig.	29	28.0	28.0	0.0	0.0	not sig.
<u>Draft #3</u>												
A	50	86.7	101.8	15.1	12.8	p<.001	45	29.1	36.5	7.4	11.0	p<.001
E	16	83.9	98.7	14.8	6.3	p<.001	-- <sup>b</sup>	--	--	--	--	--
F	69	84.3	94.2	9.9	8.7	p<.001	68	27.8	30.7	2.9	5.2	p<.001
G	36	86.2	92.5	6.3	4.3	p<.001	36	26.9	27.7	0.8	1.2	not sig.

<sup>a</sup> t-test for matched pairs.

<sup>b</sup> No data were available.

### *Qualitative Analysis*

The gain scores, statistically significant or not, do not shed any light on what ideas students tend to learn in their study of *Science: A Way of Knowing*. The curriculum developers wished to know what course content was generally successful and what content required improvement. The McNemar chi square item analysis identified specific ideas for which students showed significant increases or decreases in understanding. Table 3 summarizes the achievement for both draft #2 and #3 field trials. A significant increase or decrease in a correct response between the pretest and posttest is indicated by a "+" or "-" respectively. Although Table 3 shows no differentiation among levels of probability, it may be worth noting that for the draft #3 group: (a)  $p < .001$  for 38 items; (b)  $p < .01$  for 22 items; and (c)  $p < .05$  for 29 items. The test items in Table 3 are categorized by their content. The arbitrary classification scheme is purely a heuristic one.

Even though there is some variation in student achievement between drafts #2 and #3, the achievement of the two groups generally appears to be quite consistent. The few inconsistent areas are: tentativeness of scientific knowledge in general, the value of opinion versus fact, inquiry tactics used by scientists, assumptions, and acceptance of scientific results. (The latter two ideas did experience some improvement in draft #3.) It is interesting to compare these results with those of another study involving a unified science course (Klopfer & McCann, 1969). The consistency in all student achievement between successive years in the Klopfer and McCann study was equivalent to the consistency between successive years in those areas cited above as being "inconsistent." Thus, draft #2 and #3 groups appear to be very similar in the ideas they learned. Considering that the total gain scores varied widely among teacher groups (Table 2), it would seem that student achievement takes place within a fairly consistent cluster of items. In other words, Table 3 lists those items which generally meet with success regardless of the total *quantitative* gain of the group; that is, regardless of the teacher. Therefore, while different teachers appear to influence the quantitative gain in achievement, they influence the qualitative content of that student achievement much less for *Science: A Way of Knowing*. This is useful information for the developers of the course.

Given this fairly consistent student achievement among different teachers, the germane question becomes: *What ideas* do students learn about science and scientists? Superficially this may be answered by noticing the descriptive titles for the items which show gains during the year (Table 3). However, the question is better answered in depth by carefully examining the content of the items clustered together in Table 3.

Because the major purpose of this study is to explore the value of such information, and because space does not allow for a complete analysis here, four clusters of items have been selected for detailed analysis (aims of science, role of opinion and fact in science, "the scientific method" myth, and the role of classification in science). The four clusters were selected because they exemplify the various assets and different problems encountered in the qualitative analysis technique.

The results compiled in Table 3 indicate that students learned about the aims of science. In the past, several surveys had indicated that students confuse science with technology, especially the principal aim of each endeavour (Korth, 1969; Hurd, 1969; Mitias, 1970). In the course being evaluated, a unit entitled

TABLE 3  
SUMMARY OF ITEMS EXPERIENCING A SIGNIFICANT CHANGE  
BETWEEN PRETEST AND POSTTEST <sup>a</sup>

Topic	Number of Items		Draft #2	Draft #3
	SPI	TSAS		
<u>Aims of Science</u>	3	4	+ <sup>b</sup>	+
		1	-	-
<u>Epistemology of Science</u>				
Assumptions	4		+	+
	1		+	
	5			+
	2		-	
General Tentativeness versus Truth	1		+	+
	2		+	
		1	+	+
		1		+
Observations	2		+	+
Measurements	2		+	+
	2			+
Hypotheses	4		+	+
	1		+	
Laws	3		+	+
	2			+
	1		+	
Theories	2		+	+
	1			+
	1		-	
Models	7		+	+
	1			+
<u>Tactics of Science</u>				
"The Scientific Method" Myth	1	1	+	+
	1			+
Classification	3		+	+
	1			+
Inquiry	6		+	+
	2		+	
	2			+
	1		-	
Induction/Deduction	2		+	+
	1			+
<u>Values Within Science</u>				
Simplicity	3		+	+
Acceptance of Results	1		+	+
	4			+
	2		+	
Quantification	1		+	+
Practicality	1	2	+	+
		1	+	
Open Communication	1		+	+
		1		+
<u>Interaction of Science With Society</u>		2	+	+
		2		+
		1	+	
		1	-	
<u>Human Characteristics</u>		3		+
		3	+	

<sup>a</sup> A detailed identification of individual items is available from the author.

<sup>b</sup> A statistically significant McNemar chi square value (0.05 level of probability) is signified by the symbols "+" for an increase in correct response and "-" for a decrease in correct response.

“Technology: A Way of Knowing” distinguishes between a technologist (providing products and techniques for public use) and a scientist (satisfying his or her curiosity about the natural world). While this dichotomy is as simplistic as “pure” and “applied” science, it was nevertheless useful in bringing some order out of



TABLE 4  
ITEMS CONCERNING THE AIMS OF SCIENCE

	Items	Correct Response		% Response			$\chi^2$ <sup>b</sup>	p <
				agree	? <sup>a</sup>	disagree		
T9	The primary objective of the working scientist is to improve human welfare.	disagree	pretest posttest	65.6 39.6	4.7 0.7	29.7 59.7	33.2	.001
T22	The principal aim of science is to provide the people of the world with the means for living better lives.	disagree	pretest posttest	75.8 40.0	1.3 2.1	22.8 57.9	32.9	.001
T24	Science is primarily a method for inventing new devices.	disagree	pretest posttest	33.8 23.6	3.4 0.7	62.8 75.7	7.1	.01
S129	The primary objective of science is to develop new and improved living conveniences.	disagree	pretest posttest	56.2 31.5	3.0 0.0	40.8 68.5	26.5	.001
T34	The greatest accomplishments of science consist of the many useful products it has produced.	disagree	pretest posttest	73.6 47.3	4.1 4.1	22.3 48.6	25.9	.001
S86	One of the aims of science is to work towards more complex knowledge.	disagree	pretest posttest	75.4 62.0	5.3 0.6	19.3 37.4	16.9	.001
S94	Prediction is an important goal of scientific investigation.	agree	pretest posttest	67.5 77.8	3.6 1.8	29.0 20.5	6.6	.05
T43	The aim of science is to increase man's knowledge of the physical and biological world.	agree	pretest posttest	88.5 78.4	2.7 1.4	8.8 20.3	5.5	.05
S51	The main object of basic scientific research is the discovery of understanding rather than its practical application.	agree	pretest posttest	64.1 71.3	20.6 12.3	15.3 16.4	2.7	NS <sup>c</sup>

<sup>a</sup> ? means "do not understand the meaning"  
<sup>b</sup> McNemar chi square  
<sup>c</sup> NS means "not significant"

chaos in the minds of adolescents. Later in the course the two extremes are integrated in a less naive fashion in the context of the everyday world. (The philosophical stance that science searches for truth was rejected by the authors of *Science: A Way of Knowing*. The course basically embraces the Thomas Kuhn approach to understanding science.) The specific items listed in Table 3 under “aims of science” may be found in Table 4. The first four actually refer to a principal aim of technology. Many students seem to learn that science is *not* technology but do they learn what science is? TSAS item 43 and SPI item 51 are the only ones in both instruments which state a major aim of science. The decrease in learning reflected by item TSAS 43 and the lack of a more positive result on item SPI 51 may be explained by the result that students rigidly focussed on the curiosity aim, and therefore they do *not* translate that idea into human behaviours such as those described in item TSAS 43 and SPI 51. Therefore, draft #4 of *Science: A Way of Knowing* will have to require that students explore how the curiosity aim manifests itself in a number of ways in university and industrial laboratories. Not only does one section in the technology unit need to be rewritten, but several parts throughout the course will need modifying to reinforce the idea.

Table 3 contains a blatant example of misunderstandings apparently being learned (subsection “opinion versus fact”). A closer analysis yields useful information for the curriculum developer and evaluator (see Table 5). In order to understand how students interpret some of these items, a group not involved in the study was asked to write a justification for their responses. By reading student written responses, one finds that students tend to view SPI item 41 as simplistic and therefore reject it on that account. They fail to see its underlying sophisticated

TABLE 5  
ITEMS CONCERNING THE ROLE OF OPINION AND FACT IN SCIENCE

Items		Correct Response		% Response			$\chi^2$	p <
				agree	?	disagree		
S41	Science must start with facts and end with facts no matter what theoretical structures it builds in between.	agree	pretest	55.0	6.4	38.6	11.6	.001
			posttest	39.4	8.2	52.4		
S44	The assumptions in science are based on opinion, not fact.	disagree	pretest	50.3	2.3	47.4	9.1	.01
			posttest	65.7	1.2	33.1		
S60	Scientific conclusions should be based on facts, not opinion.	agree	pretest	84.8	0.6	14.6	6.6	.05
			posttest	74.1	0.0	25.9		
T45	The scientific investigation of human behaviour is of little value since it must involve the personal opinion of the investigator.	disagree	pretest	39.0	9.6	51.4	3.9	.05
			posttest	28.4	10.1	61.5		

meaning. On the other hand, the “correct” affirmative student responses encompass a wide range of justifications, ranging from naive to sophisticated. Quite differently, SPI items #44 and #60 (Table 5) represent an interesting problem for the developers of *Science: A Way of Knowing*. In the unit “An Artistic Way of Knowing: Perception,” students become involved with the role of basic assumptions in perception. Students study the influence that culture in general and education in particular have on these assumptions. They learn that different educational backgrounds likely lead to different perceptions or world views. Apparently this relativism encourages students to associate opinion with a world view; “world views are culturally relative thus they are in the realm of opinion, not fact,” they reason. Students consequently believe that a scientist’s world view (perceptions, assumptions, paradigm) is based on opinion. However, it seems clear that the SPI items associate “fact” with “world view” and consequently for these items “opinion” has the connotation of conjecture. This connotation is not shared by students who have adopted the relativism position inherent in the course. One can discern this relativism in item TSAS #44 (Table 5). Understandably, there is a significant increase in correct response for this item. The interesting problem which emerges from this analysis is that “negative” learning (SPI items #44 and #60) is actually “positive” learning (TSAS item #45). The definition of what is negative and what is positive relies on one’s philosophical stance (Lucas, 1975). Curriculum developers must be very aware of the curriculum’s philosophical stance so that they can properly interpret student achievement in terms of that stance (Campbell, 1971). In the case of *Science: A Way of Knowing*, the “negative” test results for SPI items #44 and #60 do not seem to warrant any specific changes in the curriculum materials. The apparent misunderstanding over opinion versus fact turns out to be more an artifact of the test itself.

At the end of the course, students believed less in “the scientific method” as a step by step procedure which: advances knowledge, produces correct answers, and leads to successful experiments (Table 6). However, a majority of students on the posttest (55%) still expressed the belief that there was such a method. The developers may not be satisfied with the dramatic positive jump from 25% to 43% in correct response. If they are not, then many places in the student’s manual will need to be rewritten in a way that challenges the application of “the scientific method.” Already in the course there are questions and discussions dealing with the topic. More would seem justified.

TABLE 6  
ITEMS CONCERNING “THE SCIENTIFIC METHOD” MYTH

Items	Correct Response		% Response			$\chi^2$	p <
			agree	?	disagree		
T3	disagree	pretest	69.6	5.4	25.0	10.9	.001
		posttest	55.4	2.0	42.6		
S84	disagree	pretest	32.4	2.4	65.3	7.1	.01
		posttest	22.9	0.6	76.5		
S124	disagree	pretest	20.4	2.4	77.2	9.8	.01
		posttest	11.2	0.6	88.2		

Students made major gains in their appreciation of classification schemes in science (Table 7). It is interesting to note that SPI item 61 is for the most part not understood. From reading students’ written justifications, it is apparent that the word “inherent” is the major reason for this. The developers of *Science: A Way of Knowing* apparently need not make any alternations regarding the nature of classification schemes in science.

TABLE 7  
ITEMS CONCERNING THE ROLE OF CLASSIFICATION IN SCIENCE

Items	Correct Response		% Response			$\chi^2$	p <
			agree	?	disagree		
S52	agree	pretest	47.4	43.3	9.4	22.5	.001
		posttest	73.1	15.8	11.1		
S61	disagree	pretest	27.2	62.7	10.1	7.8	.01
		posttest	30.4	49.7	19.9		
S78	agree	pretest	63.5	27.1	9.4	26.7	.001
		posttest	87.1	7.1	5.9		
S83	agree	pretest	81.2	7.6	11.2	6.8	.01
		posttest	90.0	1.8	8.2		

Discussion

The formative evaluation process can be greatly enhanced by using qualitative data. As seen in the few sample analyses above, the developers of *Science: A Way of Knowing* received information which had a direct application to revising their materials and activities. Qualitative data are substantial valid sources of feedback in formative evaluation.<sup>1</sup>

The usefulness of qualitative data can be clarified by noting what information an evaluator receives when his work goes beyond the usual quantitative analysis.

The quantitative data generated in this study reflected a significant increase in students’ understanding about the processes of science. However, the *qualitative* data clearly indicated what these understandings were: assumptions underlying science; the nature of observations, measurements, hypotheses, laws, theories, models, and classification schemes; the mythology of “the scientific method”; the notion of induction and deduction; and values related to simplicity, acceptance of results, quantification, and practicality of scientific results.



Unlike the quantitative analysis, the qualitative analysis was able to reveal some apparent misunderstandings inherent in the course materials. In most cases, however, the philosophical positions between the test items and the course objectives were different. These apparent misunderstandings can be dealt with in the appropriate way by the course developers.

The quantitative analysis clearly suggested that *Science: A Way of Knowing* does not generally address itself to ideas about the social aspects of science. However, the *qualitative* data delineated ideas which were generally learned: the distinction between science and technology, the relationships between the two, the importance of open communication to the advancement of science, the tentativeness of scientific knowledge, some human characteristics of scientists, and the role of science in helping to solve social problems. Knowing this, the developers can make changes that simply emphasize these ideas, thereby affecting the quantitative achievement of students. At the same time, completely new material that may affect the qualitative achievement would not seem necessary for these same ideas. On the other hand, if there are item clusters which show no significant improvement, and if the item clusters relate to a desired goal of the course, then the developers may need to construct new activities and materials.

The quantitative data indicated that student achievement varied with the teacher. However the qualitative data showed a high degree of consistency among the teacher groups in terms of what content was learned. This finding is relevant to inservice training for *Science: A Way of Knowing*. Once teachers have a consistent philosophical posture with respect to the major content of the course (as these teachers seem to have had), some will need to develop teaching techniques in order to emphasize this content.

The quantitative data gave no indication of the knowledge deficiencies students had at the end of the course. On the other hand, the qualitative data explicitly showed that there was still room for improvement in such areas as: the analysis of an experiment in terms of its control and its variables, the tentativeness of scientific laws, a lingering belief in "the scientific method," and the unwarranted faith in scientific beliefs and attitudes outside the realm of scientific inquiry. These items experienced less than 50% correct student response.

By noting the proportion of students who thought they did not understand the meaning of an item, the evaluators were able to pinpoint specific reasons why certain items were not correctly answered. The wording of some SPI items turned out to be particularly difficult (not understood by more than 25% of the students); for example: #43 ("experimental control"), #57 ("control of variable"), #61 ("inherent"), #120 ("tentative"), and #132 ("degree of estimate"). This feedback illustrates a further reward from using qualitative data. An evaluator ought to be able to distinguish between students' misunderstanding an idea and their not understanding an item's wording.

Certainly the content of the SPI and TSAS does not encompass all the major objectives of the course. The ones excluded must be evaluated in some other way (Champagne & Klopfer, 1974). On the other hand, there is likely to be content in the SPI and TSAS which falls outside the objectives of *Science: A Way of Knowing*.<sup>2</sup> Nevertheless, it seems advisable for a curriculum developer to assess the impact of a course beyond the parameters of its specific objectives (Aikenhead, 1974a). Klopfer (1969) clarified this issue when he described the *effectiveness* of a course in terms of how closely it met its objectives, while its *effect* embraced a

wider range of outcome. For this reason, it was worth analyzing all the items on the SPI and TSAS, whether or not the developers perceived them as relevant to the course. If the course is having positive or negative effects unanticipated by the developers, they should be aware of these.

In a number of instances, the SPI and TSAS seemed to assess a student's *sophistication* in knowledge about science and scientists. This phenomenon seems important to recognize when using qualitative data. For the purpose of the present discussion, one can simply assume that from *ignorance or confusion* a student develops a *basic* understanding and then from there a *sophisticated* understanding. Before she can become sophisticated in an area of knowledge, she must first have a basic, perhaps simplified, understanding. Many of the Saskatchewan tenth grade students harboured misunderstandings about science and scientists which had to be improved to simple understanding before sophistication could be achieved. A number of items from the SPI and TSAS would seem to reflect an improvement from simple understanding to sophisticated understanding, while other items reflect an improvement from misunderstanding to simple understanding. Therefore over a one-year period, there may be a little improvement on a number of items of sophistication because students are involved with correcting misunderstandings. This might explain the results concerning the "aims of science." Students tended to correct a misunderstanding (the confusion of science with technology) but they were not able to think of the principal aims of science in the sophisticated way anticipated by certain test items (Table 4). Conversely, when students have a *basic* understanding of some idea, one would not expect to see much improvement on relevant items of basic understanding. However one would hope to notice significant improvement on items assessing sophistication in that same idea.

While qualitative data provide a curriculum developer with feedback having much less ambiguity than quantitative data, qualitative data naturally have limitations. Because students bring their subjective interpretation to each item, their response on a Likert five-point scale has inherent subjectivity. In a number of instances, the developers of *Science: A Way of Knowing* could only interpret student responses by having students write a justification for their responses. These written responses revealed a range of misunderstandings as well as a range of understandings — from simple to sophisticated. This written information gave some insight into the subjectivity of "objective" tests. It also lent considerable insight into the impact of the course. Because students related to the course content in idiosyncratic ways, their written responses often exposed these idiosyncracies and allowed the reader some opportunity to assess the kind of impact the instruction had had. *Future work in formative evaluation might devise some fairly reliable techniques for achieving an in-depth analysis of student learning.*

The qualitative analysis procedure proved to be a valuable tool in the repertoire of techniques used in formative evaluation. The role of qualitative data appears to be more extensive in formative evaluation than it is in summative evaluation, a prediction which prompted this study (Aikenhead, 1974b).

#### Notes

1. Other sources include: student achievement on tests constructed specifically for the course content; student achievement in activities directly related to those of the course; student achievement in analyzing science and technology content of the mass media; parental reaction to their child's work; the effect on student achievement in subsequent classes, not only



- science classes, but other subject areas as well; the effect on the school from having an interdisciplinary offering. For a complete analysis of the formative evaluation process, see Champagne and Klopfer, 1974.
2. The TSAS, for instance, contains a number of items related to a scientist's ethical responsibility to society. *Science: A Way of Knowing* does not focus on this issue; it is left to a subsequent course where students are more mature to handle this content.

### References

- Aikenhead, G. S. The measurement of knowledge about science and scientists: An investigation into the development of instruments for formative evaluation. Unpublished doctoral thesis, Harvard University, Cambridge, Mass., 1972.
- Aikenhead, G. S. The measurement of high school students' knowledge about science and scientists. *Science Education*, 1973, 57, 539-549.
- Aikenhead, G. S. Course evaluation. I: A new methodology for test construction. *Journal of Research in Science Teaching*, 1974, 11, 17-22. (a)
- Aikenhead, G. S. Course evaluation. II: Interpretation of student performance on evaluative tests. *Journal of Research in Science Teaching*, 1974, 11, 23-30. (b)
- Aikenhead, G. S. Learning about the scientific enterprise in Saskatchewan high school chemistry classes: An evaluation. *Saskatchewan Journal of Educational Research and Development*, 1975, 5, 4-29. (a)
- Aikenhead, G. S. A new kind of science experience in the high school. A paper presented to the 23rd National Convention of NSTA, Los Angeles, March, 1975. (b)
- Aikenhead, G. S. New developments in Saskatchewan' chemistry curriculum. A paper presented to the 17th Annual Conference of the Science Council A.T.A., Banff, Alberta, Canada, October 9, 1977.
- Aikenhead, G. S., & Fleming, R. W. *Science: A way of knowing*. Saskatoon, Sask.: Department of Curriculum Studies, University of Saskatchewan, 1973.
- Campbell, D. C. Epistemological posture as an intellectual variable. Ontario Institute for Studies in Education: Department of Curriculum, 1971 (mimeographed).
- Champagne, A. B., & Klopfer, L. E. Formative evaluation in science curriculum development. *Journal of Research in Science Teaching*, 1974, 11, 185-203.
- Cooley, W. W., & Klopfer, L. E. The evaluation of specific educational innovations. *Journal of Research in Science Teaching*, 1963, 1, 73-80.
- Hurd, P. *New directions in teaching secondary school science*. Chicago: Rand McNally & Co., 1969.
- Klopfer, L. E. Effectiveness and effects of ESSP astronomy materials. *Journal of Research in Science Teaching*, 1969, 6, 64-75.
- Klopfer, L. E., & McCann, D. C. Evaluation in unified science. *Science Education*, 1969, 53, 155-164.
- Korth, W. W. The use of the history of science to promote student understanding of the social aspects of science. Unpublished doctoral thesis, Stanford University, 1968.
- Korth, W. W. Test every senior project: Understanding the social aspects of science. A paper presented at the 42nd meeting of the National Association for Research in Science Teaching, Pasadena, California, February 8, 1969.
- Lucas, A. M. Hidden assumptions in measures of "knowledge about science and scientists." *Science Education*, 1975, 59, 481-485.
- MacKay, L. D. Development of understanding about the nature of science. *Journal of Research in Science Teaching*, 1971, 8, 57-66.
- McNemar, Q. *Psychological statistics* (4th ed.). New York: John Wiley, 1969.
- Mitias, R. G. E. Concepts of science and scientists among college students. *Journal of Research in Science Teaching*, 1970, 7, 135-140.
- Tamir, P. Understanding the processes of science by students exposed to different science curricula in Israel. *Journal of Research in Science Teaching*, 1972, 9, 239-245.
- Welch, W. W. *Welch Science Process Inventory, Form D*. Minneapolis, Minn.: Dr. Wayne Welch, 1969.



## BOOK REVIEWS

RESEARCH INTO TEACHING METHODS IN HIGHER EDUCATION, 4th Edition. By Ruth M. Beard, Donald A. Bligh and Alan G. Harding. Guildford, Surrey: Society for Research into Higher Education Ltd., 1978, 150 pp. (ISBN 0-900868-58-9)

The fourth edition of this monograph is an updated and somewhat enlarged version of one which appeared first in 1967. The Society for Research into Higher Education serves universities, polytechnics and colleges of education in Great Britain; this monograph is one part of that program of service focussed on British problems. One notable quotation from the Introduction provides a partial background:

Until comparatively recently all changes in university teaching were due to outstanding innovators in the universities, or followed on recommendations of committees and professional bodies. Few of these have been directly influenced by findings in the psychology of learning or experiments into the effectiveness of teaching methods; they were based almost exclusively on teachers' views as to how the subject should develop, their experience of learning and teaching and knowledge of methods used elsewhere. (p. 2)

For people interested in the improvement of instruction in colleges, technical institutes and universities, the Beard et al. work is a useful review. However, because of its deliberate focus upon British sources to the exclusion of North American and other sources, it suffers from a special kind of flatness of perspective.

A list of chapter headings may provide potential readers with some clues about the relevance of the monograph to their own research and teaching interests:

Chapter 1: Aims and Objectives; Chapter 2: Economy and Efficiency; Chapter 3: Recall and Retention of Information (including fairly lengthy sections on: programmed learning, the lecture, and audio-visual aids); Chapter 4: Skills and Abilities (including: laboratory skills, study skills, oral skills and group discussion, and higher mental skills); Chapter 5: Teaching for Change of Attitudes; Chapter 6: Evaluation of Students, Teachers, and Teaching Methods.

The bibliography contains over five hundred entries and is, in itself, a valuable set of information. Perhaps the monograph would benefit from the newly-developing meta-analytic approaches to research summaries which Glass (Glass, Gene V., "Primary, Secondary, and Meta-Analysis of Research" in *Educational Researcher*, November 1976) and some of his colleagues have been working on; for the moment, there are still no answers to questions about generalization, a theory of teaching, or, even, future directions for research. The authors have not even attempted a "box-score" of the type that Barak Rosenshine (Rosenhine, B., *Teaching behaviors and student achievement*. National

Foundation for Educational Research in England and Wales, 1971) has made famous in the world of research in K-12 settings. Although “bottom-lines” are not discernible, the book undoubtedly serves its own constituents and is of potential interest to researchers and teachers in post-secondary institutions elsewhere.

A. MacKay  
*Centre for Research in Teaching*  
*The University of Alberta*

GLADLY WOULD HE TEACH. *By John W. Chalmers.* Edmonton, Alberta: Alberta Teachers' Association Educational Trust, 1978, 197 pp., \$5.00.

The story of Milton Ezra LaZerte is almost synonymous with the entire history of education in Alberta. Described as an “educational dynamo,” he served as a public school teacher, principal, school inspector, teachers' association president, university instructor, dean of education, trustee, chairman, and alderman. He was a member of most local, provincial and national educational bodies and, as the author of this biography says, “tended to become president of every organization he joined.”

LaZerte died on February 2, 1975, at the age of 90. The Education Society of Edmonton marked its Golden Jubilee by commissioning this biography to honour “one of its most distinguished members and one of Canada's most outstanding educators.” John Chalmers, the author, is a University of Alberta education professor emeritus who has served in many fields of education in the province and who has written several books.

LaZerte, with a B.A. in mathematics and physics, began his teaching career in 1909 in Ontario, the following year moving to Calgary, Alberta, where he took his teacher training. (Those were the days when the normal school teacher training course was just four months, with practice teaching limited to an almost irrelevant scant eight lessons in a city classroom.) Until 1917, when he joined the army, he was in turn teacher, principal, and school inspector in various Alberta centres. Following the war, in 1919 he returned to the field of education in Alberta. In 1920, he enrolled as a student at the University of Chicago which was then a hotbed of the new “progressive education” of John Dewey. His studies there, particularly under psychologist Charles Judd, profoundly affected his future, convincing him that the education of students in Alberta was not education at all. He became committed to the idea of professionalization of teachers and looked ahead to the time when all teachers would be university trained.

The dream of a full-fledged school of education at the University of Alberta became a reality in 1929, when it opened with a student body of seven with LaZerte as head. The next few years were an uphill struggle for the school of education, partly because its training term was longer than that of the normal schools (in those Depression days, it was usually essential to get a paying job as quickly as possible). The 1930's, however, brought many educational changes and by 1940, the school of education was transformed into a College of Education with LaZerte as first Dean.

Despite nominal retirement at 65, LaZerte continued his creative professional career for another incredible 25 years as an academic activist. In addition to being a member of many local educational groups, he became a prime mover in Canadian educational associations, the Canada-U.S. Committee of Education, and also was a Canadian delegate to the international UNESCO Conference in 1950.



Throughout his career, he maintained his early interest in mathematics, and continually pressed for educational research. He wrote texts and articles, and invented a “problem cylinder” which predated Cusinaire and other mechanical and audiovisual teaching methods. Inclined to be a bit stiff outwardly, always wearing his black derby hat (even on a duck hunt!), LaZerte nevertheless had a tremendous influence on Canadian education. Chalmers, like most writers of commissioned biographies, has undoubtedly stressed the positive side of his subject’s life. LaZerte was, says Chalmers,

A man for all seasons: peace and war, economic prosperity and depression, educational progress and stasis. But most of all he was a man for the future. . . . He prophesied, warned, exhorted, planned and built for a better tomorrow.

Wilfrid J. Bennett  
*West Vancouver School Trustee*

THE PROCESS OF THINKING. *By Marc Belth.* New York: David McKay, 1977, 234 pp.

Marc Belth, in his “Preface,” lists an ambitious set of questions as the basis of the book’s explorations. Predictably, only three at the most are given detailed consideration. These are: (i) “What are the instruments of the thinking process?” (ii) “What can be said about the thinking process so that it can be learned?” (iii) “What are the impediments to thinking, to learning, to learning how to think?” (p. xi). Belth’s motivation seems to be derived from the belief that even with our best intentions, the world’s problems cannot be solved without adequate thinking, and since the characteristics of adequate thinking can be taught and learned, it is important to attempt to focus on these characteristics with as much clarity as possible. The main burden of the book is then to identify model-making and metaphor and analogy creation as the essential characteristics of the thinking process in a number of fields — poetry, history, and the physical and social sciences. Many common and problematic assumptions are tackled en route. Among them are the belief that improvement of thought relies upon the elimination of metaphor, the belief that we can have direct (metaphor-free) knowing, and the belief that thought, being private, cannot be an object of analytic examination.

Belth, therefore, is not simply describing or explaining the thinking process as such. He is concerned with good thinking. And since he defines thinking as “the process of analogizing” (p. 5), “good” thinking then becomes the creation and employment of analogies, metaphors, and models which are “better” than those previously employed (p. 76). In short, to the usual assumption that improved thought is at least truer thought, Belth wishes to add the detail that truer thought means truer analogies. Crucial, then, must be his account of the way in which we can tell whether one analogy is truer than another. However, this explicitly epistemological discussion, in my opinion, could be more thoroughly developed. The prevailing assumption seems to be that the meaning of truth is correspondence (see e.g., p. 64), though this is not discussed in the context of other possibilities. So far as the justification of the acceptance of a model as true, or truer than another, is concerned, the discussion in the context of history and social science for the most part assumes the obvious correctness of the theory of verification by observation. This assumption has produced some of the hottest and most sophisticated debate in recent philosophy of science. Failure to assess this debate and draw conclusions which avoid the crucial stumbling blocks detracts significantly from the value of the book. Nowhere do I find even the attempt to construct a clear and decisive



argument settling once and for all the question whether producing “better” models amounts to anything more than creating more interesting fantasies about a hypothetical “world.” There is at least one suggestion of a pragmatic criterion (p. 154), but there is no account of why we should interpret truth in pragmatic terms or why, alternatively, we should give up truth and opt for usefulness which we can gauge. I am bothered in this respect by the question: Is it true that a proposed model is more useful? Without such an explicit construction of a theory of verification, the author’s challenge of entities outside the scope of perceptual observation and physicalistic theory (labelling them, for example, “mysterious metaphysical existences” [p. 207]) is questionable.

Requiring further discussion is Belth’s adopted meaning of “thinking.” We are told that “the only time that models are not in use are those occasions when no thinking occurs” (p. 60). If “thinking” has been defined as model-making (as it has), this must be circular and quite uninformative, that is, unless there is genuinely and obviously no reason to term as “thinking” anything which does not involve creating and (critically) using models. But there is, I think, an excellent reason to widen the scope of the definition of thinking. For where the emphasis is on good thinking and on improving thinking, there is an appeal to the crucial mental ability of “perfecting” a model by using it “reflectively” and “critically.” This ability, however, must surely be beyond the use of models, for the mere substitution of one model for another could not by itself constitute a better description or explanation. The conclusion would seem to be that thinking improves on account of that aspect of its activity which involves direct (model-free) insight into the similarity of form between model and events investigated with the model. This very ability which Belth would seem to require in his theory to render intelligible his concept of improvement of thought, he explicitly denies. He inconsistently assumes such insight in the concept of model-perfecting, and when, for example, he uses the concept of “raw” data (p. 198); his stated position, however, is that we cannot “look at nature directly” (p. 61). And because model-free insight enters the discussion only tacitly, a very important topic in any consideration of thinking is ignored — the great and genuine mystery of how, as subjects, we can have an objective grasp on the character of the world. The only objectivity he offers without adequate argument is that of the public examinability of models, and that is unsatisfactory since the expression of any cultural belief of universal scope may be publicly examined and an indefinite amount of evidential support claimed, as Polanyi and Feyerabend, for example, have shown (Polanyi, 1962, ch. 9; Feyerabend, 1961).

Bearing in mind the unacceptable narrowing of the meaning of ‘thinking,’ and the actual scale of achievement in the book, I would think that a more accurate title would be: “The Role of Model, Metaphor and Analogy in Thinking.” That models are indispensable vehicles in the development of thought is a thesis amply defended by Belth, but no satisfactory case is made for the identification of thought with the process of model invention and use. Without this, the title as it stands is misleading.

If, therefore, the concept of model-improvement involving truth value is to stand and if that logically requires some form of direct insight into the nature of things, two interesting problems arise. Firstly, if we can know the world directly, why do we need models, metaphors, and analogies? Secondly, if we have no direct knowledge, how can we check the truth value of our models? It could, I think, have been extremely illuminating about the process of thought if Belth had slightly, but significantly, reoriented his wealth of material to address these questions.

Belth makes it clear right at the start that his concern is to encourage the

attempt to understand thought in such a way as to liberate its fullest possible effectiveness in the actual world of present existence (pp. xviii-xxi). With the soundness of this intent it would be hard to disagree, but there is some doubt about the wisdom of Belth's own approach to the solution of this problem of relevance and of concreteness. I suspect that for the average graduate or upper undergraduate student, there would be adequate concrete illustration only in the chapter on poetic thought. For most, the mass of abstract generalizations could easily be lacking significant meaning. Even in the chapter on scientific thinking, the attempt to work concretely in terms of the example of Melzack's theory of the perception of pain is rendered largely ineffective, for at the very point where the effectiveness of the example is to be realized — the explicit connection of the details of the example with the list of abstract elements of a scientific model — we are merely told that "each of these elements can be recognized in the illustration" (p. 194). But can they? I very much doubt it, except by readers whose sophistication in the understanding of contemporary model theory is close to Belth's own.

There is awareness in the presentation of ideas that if interest in understanding thinking is to occur, the vagueness, paradoxicality, and apparent groundlessness of some of our most pervasive assumptions about the process must be exposed so sharply that we feel we cannot proceed without a resolution of the problems. The supreme educational art of the provocateur is required here, of which Socrates' dialectic is a fine example. For the most part, however, Belth drops his controversial theses far too soon, settling for what is almost a mere listing of alternatives before moving on. The effect is more that of an amassing of information. The one sustained and exhaustively illustrated argument is that all thinking involves models, metaphors, and analogies, but the lack of a clear sense of contrast reduces its effectiveness. For we are never shown unmistakably enough through quotational illustration that anyone seriously doubts the truth of this thesis nor what, through a relentless drawing of unavoidable implications, the unwelcome consequences of believing the reverse might be. Much is suggested, hinted at, and implied in this respect, but this is not enough to accomplish the special engagement of interest at which Belth aims.

As previously argued, the main theme itself — concerning the impossibility of thought without model, metaphor, and analogy — though exhaustively illustrated, is inadequately demonstrated on crucial philosophical points. Too few of the questions raised receive adequate treatment. Unfortunately, the introduction of too many questions hinders the author's fine intention and reduces the effectiveness of his obviously impressive background scholarship.

Foster N. Walker  
*Department of Educational Foundations*  
*The University of Alberta*

#### *References*

- Feyerabend, P. K. *Knowledge without foundations*. Ohio: Oberlin College, 1961 (mimeographed).  
Polanyi, M. *Personal knowledge*. Chicago: University of Chicago Press, 1962.



## LIST OF BOOKS RECEIVED

- Bassett, G. W., Watts, B. H., & Nurcombe, B. *Individual differences*. Winchester, Mass.: Allen & Unwin Inc., 1978, 284 pp., \$23.75 (cloth), \$12.95 (paper).
- Beard, R. M., Bligh, D. A., & Harding, A. G. *Research into teaching methods in higher education*. Surrey, England: Society for Research into Higher Education, 1978, 150 pp.
- Beck, C., et al. *Final report 1967-77: The moral education project (Year 5)*. Toronto: Ontario Institute for Studies in Education, 1979, \$3.00.
- Belth, M. *The process of thinking*. New York: David McKay, 1977, 234 pp.
- Berkeley, H., Gaffield, C., & West, W. G. (Eds.) *Children's rights: Legal and educational issues*, Toronto: Ontario Institute for Studies in Education, 1978, 177 pp.
- Billing, D. *Course design and student learning: Papers presented at the Society's thirteenth annual conference, 1977*. Surrey, England: Society for Research into Higher Education, 1978, 167 pp.
- Blaug, M. *Economics of education: A selected annotated bibliography (3rd ed.)*. Elmsford, N.Y.: Pergamon Press Ltd., 1978, 428 pp.
- Brown, E. K. *Foundations of educational evaluation*. Ardmore, PA: Dorrance & Company, Inc., 1978, 192 pp., \$10.00.
- Canadian Education Association. *Educational research and policy formation*. CEA/CERA Research Conference, 1977. Toronto, 1978, 128 pp., \$5.00.
- Canadian Education Association. *Requirements for teaching certificates in Canada*. Toronto, 1978, \$2.50.
- Carson, R. B., & Friesen, J. W. *Teacher participation: A second look*. Washington: University Press of America, 1978, 103 pp., \$6.75.
- Chalmers, J. W. *Gladly would he teach*. Edmonton: Alberta Teachers' Association Educational Trust, 1978, 197 pp., \$5.00.
- Clough, D. B., & Clough, B. M. *A handbook of effective techniques for teacher aides*. Springfield, Ill.: Charles C. Thomas, 1978, 189 pp., \$10.75.
- Collier, G. (Ed.) *Evaluating the new B.Ed.* Surrey, England: Society for Research into Higher Education, 1978, 120 pp.
- Cremin, L. A. *Traditions of American education*. Don Mills, Ontario: Fitzhenry & Whiteside Limited, 1977, 172 pp., \$5.50.
- Downey, M., & Kelly, A. V. *Moral education: Theory and practice*. New York: Harper & Row, 1978, 226 pp., \$7.50.
- Fowler, W. *Day care and its effects on early development*. Toronto: Ontario Institute for Studies in Education, 1978, 107 pp., \$7.75.
- Fowler, W. *Guides to early day care and teaching (Supplement to Day care and its effects on early development)*. Toronto: Ontario Institute for Studies in Education, 1978, 83 pp., \$3.50.
- Gagliardi, R. *The mathematics of the energy crisis*. Westmont, NJ: Intergalactic Publishing Company, 1978, 72 pp.
- Gayfer, M. *An overview of Canadian education: Second edition*. Toronto: Canadian Education Association, 1978, 43 pp., \$2.50.
- Gayfer, M. *Canadian approaches to school health education and services*. Report of a CEA survey. Toronto: Canadian Education Association, 1978, 64 pp., \$3.00.



- Glenn, J. A. (Ed.) *The third R: Towards a numerate society*. Don Mills, Ontario: Fitzhenry & Whiteside Limited, 1978, 140 pp., \$6.95.
- Hassam, A. S. *The learning process: How to pass examinations*. Hicksville, N.Y.: Exposition Press, Inc., 1978, 79 pp., \$5.50.
- Hétu, J.-C. *Stratégie d'enseignement des nombres entiers naturels*. Montréal: Les Presses de l'Université de Montréal, 1978, 150 pp., \$12.75.
- Hillerich, R. L. *A writing vocabulary of elementary children*. Springfield, Ill.: Charles C. Thomas, 1978, 295 pp., \$13.75.
- Hodgetts, A. B., & Gallagher, P. *Teaching Canada for the '80s*. Toronto: Ontario Institute for Studies in Education, 1978, 136 pp., \$5.95.
- Kelly, A. V. *The curriculum: Theory and practice*. New York: Harper & Row, 1977, 202 pp., \$6.35.
- Kurtz, A. K., & Mayo, S. T. *Statistical methods in education and psychology*. New York: Springer-Verlag New York Inc., 1978, 540 pp., \$12.00.
- Lewy, A. (Ed.) *Handbook of curriculum evaluation*. Toronto: Ontario Institute for Studies in Education, 1977.
- Mussen, P. H., Conger, J. J., & Kagan, J. *Child development and personality. Fifth edition*. New York: Harper & Row, 1979, 577 pp., \$15.95.
- Nyberg, V. R., & Lee, B. *Evaluating academic achievement in the last three years of secondary school in Canada*. Toronto: Canadian Education Association, 1978, 60 pp., \$3.00.
- Prokos, G. *This is about attendance counselling*. Toronto: Canadian Education Association, 1978, 23 pp., \$1.50.
- Ravitch, D. *The revisionists revised*. New York: Basic Books, 1978, 194 pp., \$11.65.
- Richards, J. *Classroom language: What sort?* Winchester, Mass.: Allen & Unwin Inc., 1978, 154 pp., \$16.25 (cloth), \$7.50 (paper).
- Ricker, E. (Ed.) *Education and development in Atlantic Canada*. Halifax, N.S.: Dalhousie University, 1978, 381 pp., \$8.00.
- Rosier, M. J. *Early school leavers in Australia*. Stockholm, Sweden: Almquist & Wiksell International, 1978, 198 pp.
- Sawyer, D. C. (Ed.) *Directory of education studies in Canada, 1976-1977*. Toronto: Canadian Education Association, 1978, 314 pp., \$6.00.
- Swain, M., & Harley, B. (Eds.) *Working papers on bilingualism, Issue No. 16*. Toronto: Ontario Institute for Studies in Education, 1978, 119 pp.
- Swain, M., & Harley, B. (Eds.) *Working papers on bilingualism, Issue No. 17*. Toronto: Ontario Institute for Studies in Education, 1979, 198 pp.
- Weizman, R., Brown, R., Levinson, P. J., & Taylor, P. A. (Eds.) *Piagetian theory and the helping professions*. Los Angeles: University of Southern California, 1978, 432 pp., \$12.00.

# PUBLICATIONS

FACULTY OF EDUCATION,  
UNIVERSITY OF ALBERTA  
EDMONTON T6G 2G5

## **The Alberta Journal of Educational Research**

**AJER** is a quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education. Published in March, June, September and December; subscription is \$8 per year. Address communications to the Editor, *AJER*, 732 Education South.

## **Elements**

**Elements** endeavours to help educators translate theory and findings from current research into classroom practice. Each issue focuses on a specific subject area, topic or theme. Published monthly September through April; subscription is \$4 per year. Address communications to the Editor, *Elements*, Department of Elementary Education, 539 Education South.

## **Indian-Ed**

**Indian-Ed** publishes articles, presents abstracts of research and reports news items related to the education of native peoples. Published four times per year; subscription is \$3. Address communications to the Editor, *Indian-Ed*, Department of Educational Foundations, 5-109 Education North.

## **The Canadian Administrator**

The **CA** serves administrators at all levels of educational organizations. Articles selected for publication are intended to familiarize practitioners with the findings and implications of research, and with analyses of current policies and practices. Published eight times per year October through May; subscription is \$5 per year. Address communications to Editor, *CA*, Department of Educational Administration, 7-104 Education North.

# The Canadian Journal of Sociology Cahiers canadiens de sociologie

---

## Editors

Richard Ericson  
Andrew Harrell  
Lyle Larson  
Nico Stehr

## Editorial Board

Gerald Fortin  
Edward Gross  
Stanley Lieberman  
Mildred Schwartz  
Dennis Wrong

## Associate Editors

T.R. Balakrishnan—University of Western Ontario  
Wallace Clement—McMaster University  
James Curtis—University of Waterloo  
Leo Driedger—University of Manitoba  
Larry Felt—Memorial University  
Dennis Forcese—Carleton University  
Martha Foschi—University of British Columbia  
Herbert Gamberg—Dalhousie University  
Timothy Hartnagel—University of Alberta  
R. Alan Hedley—University of Victoria  
Harry H. Hiller—University of Calgary  
John D. Jackson—Concordia University  
J. Paul Johnston—University of Alberta  
Ronald D. Lambert—University of Waterloo  
Eugen Lupri—University of Calgary  
William Michelson—University of Toronto  
James C. Moore—York University  
James Richardson—University of New Brunswick  
Brigitte Schroeder-Gudehus—Université de Montréal  
Dorothy Smith—Ontario Institute for Studies in Education  
Malcolm Spector—McGill University  
Lorne Tepperman—University of Toronto  
Roy Turner—University of British Columbia  
Kenneth Westhues—University of Waterloo

*The Canadian Journal of Sociology* is published at  
The University of Alberta. Individual subscriptions are  
\$18 per year. Subscriptions from institutions and libraries are  
\$36 per year. Requests concerning manuscripts,  
advertising and subscriptions should be sent to  
The Canadian Journal of Sociology, Department of Sociology,  
The University of Alberta, Edmonton, Alberta, Canada T6G 2H4.

Some recent and  
forthcoming articles from the  
*quarterly CJS*: **Bernd Baldus**—  
Social control in capitalist  
societies; **Alan Blum and Peter McHugh**—  
A Sense of Place; **Merlin Brinkerhoff**—  
Women who want to work in a man's world;  
**S.D. Clark**—Sociology in Canada; **Harold Fallding**—  
Mainline protestantism in Canada and the United  
States; **Harriet Friedmann and Jack Wayne**—Depend-  
ency theory; **Richard Henshel and Robert Silverman**—  
Perception and criminal process; **I.C. Jarvie**—Nationalism  
and the social sciences; **Niklas Luhmann**—Differentiation of  
society; **Neil Mackinnon and Gene Summers**—Homogeneity and role  
consensus; **Wolfgang Mommsen**—Max Weber as critic  
of Marxism; **John F. Myles and Aage Sorensen**—Elite and status attain-  
ment models; **John Rex**—Value relevance, scientific laws, and ideal types;  
**James Richardson**—Education and social mobility; **Paul Stevenson**—Frustration,  
structural blame and leftwing radicalism; **Lorne Tepperman**—Deviance as a search process.



## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.



# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXV    NUMBER 3    SEPTEMBER 1979

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    •    EDMONTON



# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

Published quarterly in March, June, September, December by the  
Faculty of Education, The University of Alberta

## CONSULTING EDITORS

J. Britton  
*University of London*

J. Calam  
*The University of British Columbia*

M. Connelly  
*The Ontario Institute for  
Studies in Education*

K. De Clerck  
*State University of Ghent*

R. N. Evans  
*University of Illinois at  
Urbana-Champaign*

R. H. Farquhar  
*University of Saskatchewan*

E. Gagné  
*University of Ottawa*

G. Harman  
*University of Melbourne, Australia*

S. Hunka  
*The University of Alberta*

J. W. G. Ivany  
*Simon Fraser University*

D. A. MacIver  
*University of New Brunswick*

L. D. Nelson  
*The University of Alberta*

W. C. Nesbit  
*Memorial University of Newfoundland*

E. Pedersen  
*McGill University*

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

## FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith  
T. P. Atkinson  
N. C. Bhattacharya  
W. T. Fagan

H. W. Hodysh  
W. W. Laing  
R. G. Martin

E. Miklos (Chairman)  
J. W. Osborne  
C. H. Preitz

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal Of Educational Research

Volume XXV, No. 3

September, 1979

## CONTENTS

Experimental Manipulation of the Recall of Narrative Material by Five-Year-Olds .....	137
<i>L. D. Travis and W. B. White</i>	
Perception of the Administrative Role of the Principal as a Factor in Student Conflict in Nigeria .....	147
<i>J. I. Nwankwo and T. O. Ohikhena</i>	
Education as a Social Control Mechanism .....	160
<i>M. K. Bacchus</i>	
Identical and Reverse Visual Pattern Recognition in Deaf Children .....	174
<i>R. Bragman and R. Hardy</i>	
Games and Game Settings for the Preschool Child .....	182
<i>W. Liedtke and G. Stott</i>	

## PERSPECTIVES

History of Childhood as History of Education: A Review of Approaches and Sources .....	192
<i>R. L. Schnell</i>	

## BOOK REVIEWS

Rejoinder to Professor L. K. Schubert's review (December 1978 issue of AJER) of <i>Thinking with the Teachable Machine</i> , by John H. Andreae, London: Academic Press Inc., 1977 .....	204
<i>A Theory of Education</i> , Joseph D. Novak .....	205
Reviewed by A. T. Pearson	
<i>Research in Education (3rd Ed.)</i> , J. W. Best, and <i>Educational Research in Classrooms and Schools</i> , L. Cohen .....	208
Reviewed by A. K. Clark	
<i>The Curriculum: Theory and Practice</i> , A. V. Kelly .....	211
Reviewed by G. S. Tomkins	

FACULTY OF EDUCATION  
*The University of Alberta*





LE ROY D. TRAVIS

and

WILLIAM B. WHITE

*The University of British Columbia*

## Experimental Manipulation of the Recall of Narrative Material by Five-Year-Olds

*The effects of four different conditions under which children of similar age and ability encountered the same narrative material were assessed by comparing the children's recall performances at 30 seconds and one week after the encounters. Results are reported from two studies.<sup>1</sup> In the first study, kindergarten children who enacted a brief story while listening to it recalled significantly more story elements, explicit information, and implicit information than children who merely listened to the story narration. In the second study, effects of two additional conditions were assessed, through comparison with those of the first study. These conditions entailed children who watched a puppet while listening to a narration, and children who covertly rehearsed the story after merely listening to it. Children who enacted the narrative recalled significantly more story elements and explicit information than all other groups. Alternative interpretations and implications are discussed. (Dr. Travis is a member of the Educational Psychology staff and Mr. White is a doctoral student at the Faculty of Education, University of British Columbia.)*

Every schoolboy knows that the coding, organization, and recall of information is basic to achievement in virtually all school subjects. Progress in any subject depends upon accurate recall of material, and few of us are unfamiliar with the frustrating experience of having memory fail us when we need it most. Not surprisingly then, teachers have a well-known and enduring interest in developing means to maximize recall and minimize failure to remember.

Recently, research on children's recall has focused on the recall of narrative material. Narrative materials are often presented in sentence or brief story form. Indications are that young children exhibit poor recall of narrative material (Brown, 1975; Paris & Upton, 1976). Poor recall is often attributed to age (Brown, 1975; Flavell, 1971), handicaps of early forms of egocentrism (Brown, 1975; Piaget, 1967), underdeveloped coding systems (Lange & Jackson, 1974; Piaget, 1976), or the absence of learned strategies (Chi, 1976; Flavell, 1970; Paris & Lindauer, 1976).

Attempts to discover means of improving recall performance have produced the following conclusions: first, recall may be increased if children are told to remember rather than merely listen (Yussen, Gagné, Garriulo & Kunen, 1974), or if they are instructed to look for ways that (stimulus) items go together (Rosner, 1971). Second, manipulation or induction of the adoption of strategies (e.g., instructing or requiring children to touch, imagine, or rehearse presented material rather than leaving them to their own devices) produces improved performance (Levin, Lesgold, Shimron & Guttman, 1975; Paris & Upton, 1976; Wilder, 1971). Third, if children are forced to code in specific ways (Trabasso & Riley, 1973) or given a code to work with (Cannizzaro, Cecchini & Musatti, 1975), performance is enhanced (within specified age ranges). The fourth, and last, suggestion is that the methods of presenting material may influence the amount recalled (Bijou, 1976; Brown, 1975; Paris & Lindauer, 1976). The foregoing conclusions, when considered together, suggest the possibility that the method of encoding information from narratives may be a source of the difficulties young children experience in attempts to recall such material (Elkind, 1971; Jablonski, 1974).

While there are several disputes over the nature of children's encoding practices, there is little evidence of controversy over the organizational immaturity and simplicity of young children's conceptual systems and their relatively unsystematic employment of these systems and strategies in coding and organizing abstract information. Accordingly, it may not be unreasonable to assume that these factors are implicated in children's difficulties in recalling verbal-narrative material. This conclusion appears to be congruent with Flavell's (1977) recent assessment of the evidence and its relation to the Genevan's endeavours.

Piaget has explicitly and repeatedly asserted that active engagement with the environment is critical to cognitive development (Flavell, 1963). In addition, Inhelder, Sinclair and Bovet (1974) have suggested that active engagement is implicated in encoding, organizing, and recalling. Following Piaget as well, Bruner (1964) has asserted that enactive representation of information ontologically precedes symbolic representation, where enactive representation is defined as "a mode of representing past events through appropriate motor response" (p. 2). It appears then, that these authors perceive enactive engagement with material as being critical to cognitive development and memory.

One can also take account of studies which have reported performance gains of various kinds, as a consequence of inducing motor activity. Levin, Ghatala, DeRose, Wilder and Norton (1975), for example, reported that motor-induced imagery appeared to constitute a highly effective discrimination learning strategy (with children in the fifth and sixth grades). Moreover, Therrien (1977) has shown how use of play improved recall of story sequence with five-year-olds. Similarly, Rubin and Pollack (1969) taught kindergarten boys to play games with objects as sounds, and in this way increased their auditory perception. Rubin and Pollack (1969) emphasized the effect of visual and visual-motor experience in the five-year-olds' ability to integrate multi-modal inputs. Likewise, Penman, Christopher and Wood (1977) reported that third grade children who learned proper use of capitalization and punctuation in the context of game involving physical body movement, exhibited performance on the criterion measures which was superior to that of the control cohorts who were instructed through presentations while they remained seated. Furthermore, other researchers have generally reported that inducement of active motor movement usually seems to



produce superior performance over other teaching strategies such as those which rely heavily on visual and oral presentation (Levin, McCabe & Bender, 1975; Paris & Lindauer, 1976; Silvern & Yawkey, 1977).

This literature, then, would appear to indicate that performance on memory measures may be improved if children are induced to encode and organize information through their own physical intercourse with the subject matter. In sum, enactive representation, or "a mode of representing . . . events through appropriate motor responses" (Bruner, 1964, p. 2), might be expected to enhance children's performances on memory measures which require retrieval of symbolic representations, since enactive representations seem to entail organization of the information into systems built on action schemes (Bruner, 1964) and young children's symbolic representations seem to be developed from their prior construction of enactive representations. This analysis suggested the following hypothesis: Children who are induced to represent the information of a narrative through their own motor actions will exhibit significantly greater recall of said information than will their cohorts who encounter the same narrative information without being induced to represent the information through motor actions, at 30 seconds and one week after one encounter with the narrative information.

The first of the two studies described below was designed to test the foregoing hypothesis and the second study was carried out to derive theoretical clarification of the results.

### *Experiment 1*

In the first study, kindergarten children were presented a brief story through one or the other of two instructional methods. The instructional methods consisted of an approach where children enacted a brief story with a puppet while listening to the narration (Group I) and a listening only condition where children merely heard the same narration of a story (Group II).

Since actions are the content of enactive representations and enactive representations are thought to form the basis of iconic and symbolic representations (Bruner, 1964), recall of information which is organized through a subject's actions may be enhanced due to such organization. It follows then, that if a five-year-old encountered narrative material and, through enactive experience with it, organized the material through his actions, his recall of that narrative material might be expected to be superior to recall of those who merely listened to a verbal narration of that material. Specifically, it was predicted that children who were presented a brief story and induced to act out that story while listening to it would have higher rate of recall of that story than children who only listened to it.

### *Method*

*Subjects.* A sample of 35 children ranging in age from 4.75 to 5.83 years were selected from three kindergartens in Vancouver, B.C. Ten children were from a private kindergarten, 16 were from a university area kindergarten, and 9 were from a university operated kindergarten. Three of these children were absent at posttests and were therefore excluded from the sample. Each child was randomly assigned to one of two groups with the restriction that the number of children from each school be approximately equal in each group.

On seven-point scales, teachers rated pupils on ability, performance,



intellectual skills, and overall intelligence. A composite score for each child was formed by combining the four scores for each child. Comparisons between groups with regard to age and ratings revealed no significant differences, [ $F(1,30) = .555$ ,  $p = .462$  for age;  $F(1,30) = .007$ ,  $p = .931$  for ratings].

*Apparatus.* Two homemade cotton sock puppets were used, one a rabbit and the other a deer. A two-part painted folding scene of a forest, a farm, and a garden was used as background for the puppets.

A tape-recorded narration of forty seconds' duration was used. The narration was taken from the reading series *Rockets* (Durr, Le Pere & Alsin, 1976, p. 72). Two alternate stories were produced from this source. In the first revised story, the word "perked" was replaced by the word "stood." In the second revision, the words "perked," "rabbit," "hopped," and "hole," were replaced by the words "stood," "deer," "jumped," and "forest."

A cassette tape recorder was used to deliver the narratives and to record the children's recall responses. All children were tested in isolation from their class by a male experimenter.

*Procedure.* As stated above, all children were randomly assigned to one of two groups. If assigned to Group I (enactive experience), the child was given a training session wherein he was taught to act out a sentence with a puppet. When proficiency was attained in the training session, the child was given the experimental treatment entailing his use of the puppet. This enactive experience condition (Group I) consisted of telling the child to listen very carefully to a short story and make a puppet act out the narration while listening to it. (Appropriate scenery props were used with this group.) If assigned to Group II (listening only experience), the child was told to listen very carefully to a short story. No scenery was used with this group.

When each child was finished listening to the story, there was a 30-second pause to allow for the removal of the puppet and scenery (for Group I) and the changing of tapes in the tape recorder (both groups). The first posttest was then administered. Posttest one (PT-1) consisted of asking the child to tell the story to the experimenter just as the child had heard it. When the child was finished responding (or remained silent for 20 seconds), the experimenter told the child to tell him everything or anything the child remembered about the story. When the child was finished (indicated either by stating he was finished or by failure to respond for 20 seconds), he was told he would have one minute to think about the story. After a 40-second pause, the child was then asked if there was anything he/she would like to add to what was already said. When the child finished responding (indicated by saying so or through a 20-second period of silence), PT-1 was terminated.

One week after each child had heard the narration, he/she was again tested on recall of the story. Each child was first reminded that he/she heard "a story" on the tape recorder and was then asked to tell the story to the experimenter. This second posttest (PT-2) was identical to PT-1 with regard to procedure.

*Analysis.* The story was categorized into Rumelhart's (Stein & Glenn, 1975) schema for analysis of narrative material. This brief story was broken down into 14 categories. Since either one does or does not recall a given instance of information, the data was treated as nominal data and recall of same was scored as one point while failure to recall was indicated by zero. If a child transposed a category but

gave the general meaning, the category was counted as correct. For example, this was the case when a child said, “He went in a garden” for the statement, “A little brown rabbit (deer) hopped (jumped) into a farmer’s garden.” Similarly, the substitution of “He looked for some food” for the statement, “It looked all around for something to eat,” was accepted. Each child received two scores for this procedure, one at PT-1 and one at PT-2. This comprised the counting for total narrative content of the brief story. Explicit information scores were tallied using a predetermined list of the explicit items in the passage. Implicit information scores were obtained in the same manner as the explicit scores from a list of predetermined implicit items.

Because the hypothesis to be tested was concerned only with results at the end of each posttest, a one-way analysis of variance was used for each measure (categories, explicit information, and implicit information), the posttest measures being the dependent variables and the experimental conditions the independent variables.

Results

The present study was designed to discover whether differences in recall performance would be produced by varying the manner in which children interacted with narrative material. Specifically, the study attempted to test the hypothesis that mean recall performance of children who enacted the narrative (Group I) while listening to it would be superior to mean recall performance of children who merely listened (Group II) to the same narrative. In order to test this hypothesis, two comparisons of three performances were made between two groups. That is, three measures of performances of two groups at 30 seconds and one week after encounters with the material were compared. As previously indicated, between-group comparisons with regard to age and rated ability

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF GROUPS I and II ON THE EXPLICIT, IMPLICIT, AND CATEGORY MEASURES AT PT-1 and PT-2

	PT-1		PT-2	
	$\bar{X}$	SD	$\bar{X}$	SD
<u>Explicit Measure</u>				
Group I	9.75*	3.624	6.5*	3.983
Group II	4.875	3.052	2.375	2.63
<u>Implicit Measure</u>				
Group I	1.125*	.957	.563*	.629
Group II	.5	.516	.125	.324
<u>Category Measure</u>				
Group I	5.125*	1.995	3.563*	2.128
Group II	2.75	1.653	1.375	1.455

<sup>a</sup>  $n = 16$  for each group

\*  $p < .05$



produced no significant differences. Accordingly, differences between group performances were not attributable to differences in age or rated ability.

The one-way analysis of variance between groups on the three performance measures (narrative content, explicit information, and implicit information), on both posttests, indicated that Group I recalled significantly more narrative material, explicit information, and implicit information than Group II. The means and standard deviations on each measure are summarized in Table 1. Group I had a higher rate of recall on all three measures at PT-1 and PT-2. Analysis of variance indicated this difference to be significant for the explicit information measure [ $F(1,30) = 16.938, p = .0003$  at PT-1;  $F(1,30) = 11.95, p = .0017$  at PT-2].

The implicit information measure was also significant at both posttests  $F(1,30) = 5.282, p = .024$  at PT-1 and at PT-2,  $F(1,30) = 5.976, p = .021$ .

Furthermore, the category measure (narrative content) also indicated significant performance differences at both posttests [ $F(1,30) = 13.437, p = .0009$  at PT-1;  $F(1,30) = 11.52, p = .002$  at PT-2].

### *Discussion*

Data from all measures at both posttests indicate that children who enacted a brief story with puppets (Group I) recalled more explicit information, implicit information, and narrative content than did those children who merely listened to it (Group II).

The results of the present study suggest that motoric organization of narrative material enhances recall (of that narrative material). However, one question concerning the results of this study should be discussed. This question concerns the possibility that rehearsal effects may confound the effects of the predictor variable (motor organization) in this study. Since the children of the enactive group (Group I) were necessarily rehearsing the substance of the story as they acted it out, and children in the alternative condition did not have to rehearse the material, one must consider the possibility that performance differences on the posttest measures are attributable to possible advantages gained by the Group I children from rehearsal (Bandura & Jeffery, 1973).

Unfortunately, the design of this study did not enable one to partial out such effect as might be attributable to rehearsal. Therefore, a second was undertaken to partial out such effects.

### *Experiment 2*

In this study, kindergarten children were presented the same brief story as in Experiment 1 through one of two instructional methods. One instructional method (Group III) consisted of an approach where children listened to a brief story, and immediately after hearing it, were told to close their eyes and try to see the story in their mind while remembering or thinking about it. The second instructional method (Group IV) consisted of children watching the experimenter enact the same story with a puppet while the children listened to the narration.

Since the data from Experiment 1 was difficult to interpret with regard to what factors (enactive organization or rehearsal) were implicated in the superior performance of the children in the enactment group (Group I), this study incorporated a condition (Group III) to provide a means of estimating the relative influence of rehearsal and that of enactive organization. Furthermore, since in the



first study there may have been a Hawthorne effect (regarding the use of puppets), Group IV was added to gain an estimate of such an effect.

Groups I and II of the first study were then compared with Groups III and IV of this study to test if varying the manner in which children interacted with narrative material would effect recall performance.

### *Method*

*Subjects.* A sample of 38 children ranging in age from 4.83 to 6.25 years were selected from two kindergartens in Vancouver, B.C. Seventeen children were from a private preschool kindergarten, and 21 children were from a public kindergarten. A loss of six subjects, due to absence at posttest times, reduced the sample to 32.

Each child was randomly assigned to one of two groups with the restriction that the number of children from each school be approximately equal in each group. As in Experiment 1, teachers rated their students' abilities. These scores were then compared to the groups in the first study. A one-way analysis of variance indicated no age difference between the four groups [ $F(3,60) = 1.817, p = .1537$ ]. Further, the ability ratings for the four groups when analyzed yielded no significant differences [ $F(3,60) = 1.565, p = .207$ ].

*Apparatus.* The same materials used in the first study were also used in this study.

*Procedure.* The same procedures as were followed in Experiment 1 were followed except for the changes in the nature of the conditions within which the narratives were encountered (described above).

*Analysis.* The children's responses were scored and analyzed as in Experiment 1. However, Groups I and II of the first study were analyzed with Groups III and IV of this study since the hypothesis to be tested was concerned with performances of groups from both studies. Therefore, a one-way analysis of variance for each measure (explicit information, implicit information, and narrative content or categories) was used, the posttest measures being the dependent variables and the experimental conditions being the independent variables.

Since the four groups did not differ significantly with regard to age and rated ability, it was assumed that the performance of the four groups at the two posttests could be interpreted with regard to treatment effects.

### *Results*

Like Experiment 1, the second study was designed to discover whether differences in recall performances would be produced by varying the manner in which children interacted with narrative material. A one-way analysis of variance between the four groups on three performance measures (narrative content, explicit information, and implicit information) from both posttests indicated some significant differences. The results are summarized in Table 2.

A one-way analysis of variance on the amount of explicit information recalled at PT-1 indicated significant differences [ $F(3,60) = 9.885, p = .0000$ ]. Similar results were obtained at PT-2 [ $F(3,60) = 7.5, p = .0002$ ]. Tukey's multiple range HSD test (Glass & Stanley, 1970) indicated that Group I performances were significantly superior to those of Groups II, III and IV ( $p = .05$ ) at PT-1 and PT-2.

Analysis of implicit information recalled at PT-1 and PT-2 did not indicate any

TABLE 2  
MEANS AND STANDARD DEVIATIONS OF GROUPS I-IV ON ALL MEASURES AT  
PT-1 and PT-2

	PT-1		PT-2	
	$\bar{X}$	SD	$\bar{X}$	SD
<u>Explicit Measure</u>				
Group I	9.75*	3.62	6.5*	3.98
Group II	4.87	3.05	2.37	2.63
Group III	3.62	2.36	1.56	1.63
Group IV	5.06	4.34	3.43	3.79
<u>Implicit Measure</u>				
Group I	1.12	.95	.56	.62
Group II	.5	.51	.12	.34
Group III	.68	.87	.56	.96
Group IV	1.12	1.08	.75	1.0
<u>Category Measure</u>				
Group I	5.12*	1.99	3.56*	2.12
Group II	2.75	1.65	1.37	1.45
Group III	1.43	1.2	.75	1.18
Group IV	2.81	2.45	1.5	1.71
<u>Total Score</u>				
Group I	16 *	5.8	10.62*	6.28
Group II	8.12	4.63	3.87	3.94
Group III	5.75	3.87	2.87	3.42
Group IV	9.0	7.46	5.68	6.04

<sup>a</sup>  $n = 16$  per group  
\*  $p < .05$

significant differences between the groups [ $F(3,60) = 2.04, p = .1173$  at PT-1;  $F(3,60) = 1.84, p = .149$  at PT-2].

Analysis of the narrative material (category measure), scored in the same manner as in Experiment 1, indicated significant between-group differences at PT-1 [ $F(3,60) = 10.587, p = .0000$ ] and at PT-2, [ $F(3,60) = 8.704, p = .0001$ ]. Using Tukey's HSD test, Group I was again found to be significantly outperforming Groups II, III and IV ( $p = .05$ ) at both posttests. Lastly, all three measures were grouped to form a fourth total score. A one-way analysis of variance gave significant between-group differences at PT-1 [ $F(3,60) = 9.87, p = .0000$ ] and at PT-2 [ $F(3,60) = 7.346, p = .003$ ]. Post-hoc measures again found Group I children to be recalling significantly more than the children in the other groups (at both posttests) as indicated by this analysis of the aggregate data.

Discussion

The present study examined the effects of four types of encounters with narrative material on young children's recall performances. The effects were examined by comparing the amount of free recall at two times after the encounters.

Data from all measures at both posttests indicate that children who enacted a brief story with puppets (Group I) recalled more narrative content and explicit

information than did those children who watched or listened to it (Groups II, III, IV). Furthermore, since analysis of age and rating data indicated that all four groups came from the same age and ability population, the results can be interpreted as indicating that the performance differences can be attributed to differences in experimental treatment. That is, no significant between-group differences were found in teacher ratings of ability or age. Accordingly, the evidence suggests that varying the manner in which children interact with narrative material affects their recall performance of that material and that motoric organization of the narrative information yields significantly higher recall rates.

However, it is suggested that this study be replicated to assess the reliability of these findings, and that another condition, in which children watch and then rehearse the story, be added to the design. The latter would allow for more precise theoretical interpretation of the results since this study suggests a need for further study of young children on covert and overt rehearsal effects. Finally, teachers of children who differ in developmental age may want to know the extent to which there are continuities or discontinuities (and when they occur) with regard to the relative effects these different sorts of encounters with narrative materials have on recall performances.

While a more complete theoretical explanation for the present results will depend upon future research which clarifies the extent to which a rehearsal effect is implicated in recall gains exhibited by children who act out stories, educators may take note that this study has shown that children who do organize information through their actions appear to recall more of such information than do those who watch or merely listen to it. Accordingly, to the extent to which the present results are reliable and where pedagogical techniques require the recall of information, superior recall performances may be expected if the teacher arranges for pupils to enactively organize the narrative material which they are expected to recall.

Note

1. The first study reported herein was described in a thesis submitted by the second author to the Faculty of Graduate Studies at the University of British Columbia in partial fulfillment of the requirement for the M.A. degree. Copies of this paper may be procured by writing to the senior author.

References

- Bandura, A., & Jeffery, R. W. Role of symbolic coding and rehearsal processes in observational learning. *Journal of Personality and Social Psychology*, 1973, 26, 122-130.
- Bijou, S. W. *Child development: The basic stage of early childhood*. New Jersey: Prentice-Hall, 1976.
- Brown, A. L. Recognition, reconstruction, and recall of narrative sequences by preoperational children. *Child Development*, 1975, 46, 156-166.
- Bruner, J. S. The course of cognitive growth. *American Psychologist*, 1964, 19, 1-15.
- Cannizzaro, L., Cecchini, M., & Mussati, T. [Operation and code in the operational development of the child: The operational proportion.] *Archivio de Psicologia Neurologia and Psichiatria*, 1975, 36(3), 271-312. (*Psychological Abstracts*, 1975, 58, No. 9962.)
- Chi, M. T. Short term memory limitations in children: Capacity or processing deficits? *Memory and Cognition*, 1976, 4(5), 559-572.
- Durr, W. K., Le Pere, J. M., & Alsin, M. L. (Eds.), *Rockets*. Boston: Houghton Mifflin, 1976, 72.
- Elkind, D. Early childhood education. A Piagetian perspective. *National Elementary Principal*, 1971, 51(1), 48-55.



- Flavell, J. H. *The developmental psychology of Jean Piaget*. Princeton, N.J.: Van Nostrand, 1963.
- Flavell, J. H. Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5). New York: Academic Press, 1970.
- Flavell, J. H. First discussant's comments: What is memory development the development of? *Human Development*, 1971, 14, 272-278.
- Flavell, J. H. *Cognitive development*. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1977.
- Glass, G. V., & Stanley, J. C. *Statistical methods in education and psychology*. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1970.
- Inhelder, B., Sinclair, H., & Bovet, M. *Learning and the development of cognition*. Cambridge: Harvard University Press, 1974.
- Jablonski, E. M. Free recall in children. *Psychological Bulletin*, 1974, 81(9), 522-539.
- Lange, G., & Jackson, P. Personal organization in children's free recall. *Child Development*, 1974, 45(1), 1060-1067.
- Levin, J. R., Ghatala, E. S., DeRose, T. M., Wilder, L., & Norton, R. W. A further comparison of imagery and vocalization strategies in children's discrimination learning. *Journal of Educational Psychology*, 1975, 67(1), 141-145.
- Levin, J. R., Lesgold, A. M., Shimron, J., & Guttman, J. *Strategies in reading comprehension 4: Pictures and young children's learning from oral prose. Technical report No. 328*. Madison, Wisconsin: University of Wisconsin, 1975. (ERIC Document Reproduction Service No. ED 112 584).
- Levin, J. R., McCabe, A. E., & Bender, B. G. A note on imagery-inducing motor activity in young children. *Child Development*, 1975, 46, 236-266.
- Paris, S. G., & Lindauer, B. K. The role of inference in children's comprehension and memory for sentences. *Cognitive Psychology*, 1976, 8(2), 217-227.
- Paris, S.G., & Upton, L. R. Children's memory for inferential relationships in prose. *Child Development*, 1976, 47(3), 660-668.
- Penman, K. A., Christopher, J. P., & Wood, G. S. Using gross motor activity to improve language arts concepts by third grade students. *Research Quarterly*, 1977, 48(1), 134-137.
- Piaget, J. *Six psychological studies*. New York: Random House, 1967.
- Piaget, J. *The grasp of consciousness*. Cambridge: Harvard University Press, 1976.
- Rosner, S. R. The effects of rehearsal and chunking instructions on children's multitrial free recall. *Journal of Experimental Child Psychology*, 1971, 11, 93-105.
- Rubin, L., & Pollack, C. Auditory perception in kindergarten children. *Journal of Special Education*, 1969, 3(2), 155-161.
- Silvern, S. B., & Yawkey, T. D. Young children's encoding of images based on motor and verbal modes of mediation. *Psychology in the Schools*, 1977, 14(1), 112-115.
- Stein, N., & Glenn, C. A developmental study of children's recall of story material. Paper presented at the Society for Research in Child Development, Denver, Colorado. St. Louis, Missouri: Washington University, 1975.
- Therrien, S. Children's play: Translating research into practice. *Early Childhood Education*, 1977, 10(2), 10-19.
- Trabasso, T., & Riley, C. A. *An information processing analysis of transitive inferences*. Princeton, N.J.: Princeton University, 1973. (ERIC Document Reproduction Service No. ED 091 043.)
- Wilder, L. Spoken rehearsal and verbal discrimination learning. *Speech Monograph*, 1971, 38, 113-120.
- Yussen, S. R., Gagné, E., Garriuolo, R., & Kunen, S. *The distinction between perceiving and memorizing in elementary school children*. Madison, Wisconsin: University of Wisconsin, 1974. (ERIC Document Reproduction Service No. ED 098 538.)

JOHN I. NWANKWO

and

TITUS O. OHIKHENA

*University of Ibadan*

## Perception of the Administrative Role of the Principal as a Factor in Student Conflict in Nigeria

*The wave of student unrest in Nigerian secondary schools has been increasingly blamed on the administration of the schools. To examine the connection between school administration and student crisis, an investigation was carried out into the extent to which the degree of congruence in the principal-student perception of the administrative role of the principal was related to the level of student conflict behaviour. Results showed that the two variables had a high negative relationship. In schools where there appeared to be a low degree of congruence, a high level of student conflict was evident, and vice versa. (Dr. Nwankwo is a lecturer at the University of Ibadan, Nigeria, and Dr. Ohikhena is Head of the Department of Educational Management, University of Ibadan.)*

The phenomenon of student conflict behaviour is as old as schooling. But the recent wave of student unrest in Nigeria appeared challenging to both school administrators and the government, and constituted a threat to the objectives of the education enterprise. Equally disturbing was the fact that some of the conventional explanations for such student group protests appeared inadequate in the light of the frequency and magnitude of recent student crises. Furthermore, conflict models such as the problem of generational gap (Green, 1952), adolescent "storm and stress" (Coleman, 1961; Davis, 1944; Elkin & Westley, 1955; Williams, 1952), the bimodal age, sex, or class distribution within the school population (Chesler & Franklin, 1968; Waller, 1965), low socioeconomic status or poor home background (Hollingshead, 1949; Oloruntinmehin, 1974), or parental insecurity (Frankenstein, 1970) could not sufficiently explain the nature or occurrence of the conflict phenomenon in the schools. In most of the conflict demonstrations, the student protesters directed their aggression toward the school administrator or some vital school property (*Daily Times*, February 5, 1975; April 16, 1975, October 22, 1975; *Daily Star*, March, 1976). Also, the parents, the mass media, and the public tended to blame recent student crises on the schools themselves (*Nigerian Chronicle*, July 5, 1974; *Daily Times*, February 10, 1975, March 25, 1976).

### *Problem*

The purpose of the study was to investigate to what extent student conflict behaviour is related to the perception of the role of the school administrator. The strong relationship between perception and behaviour has been stressed by social psychologists and behaviour scientists (Hargreaves, 1975, p. 33; Heider, 1958, p. 31; Rogers, 1969, p. 200).

Some recent studies on student protests had pointed to the centrality of the student perception of some unfavourable social or institutional issues in the student protest motives (Blackstone, Gales, Hardley, & Lewis, 1967, p. 241; Linowitz and others, 1970, p. 17). Also the importance of some degree of agreement (congruence) in role perception for the functioning of social organizations has been indicated in some recent reports on organizational behaviour (Hargreaves, 1975; Miklos, 1963; Zalkind & Costello, 1962).

Specifically, the problem was to find out:

1. whether the principal and the student perception of the administrative role of the principal were congruent and to what degree;
2. whether the degree of congruence in the principal-student perception of the principal's administrative role was related to the level of student conflict behaviour in secondary schools in the Imo State of Nigeria.

### *Theoretical Framework*

Generally, in a social organization, conflict may indicate the existence of incompatible activities in relation to information, beliefs, interests, or desires (Deutsch, 1969). One of the questions that motivated this study was the endemic occurrence of student conflict in Nigeria. It was observed that even among schools with students of comparable ages, sex, grades, background, or staff of identical composition and strength and comparable facilities or utilities, some experienced more student conflict and disruption than others. Were the schools with more frequent student conflict, then, experiencing greater differences in the members' beliefs, perceptions, expectations, interests, and desires than the less disrupted schools? To this end, we hypothesized that student conflict behaviour in schools might turn out to be students' reaction to their right or wrong perception of some differences between their perception or expectation of the school administration and what the school administrator (the principal) perceived himself or his office to be doing. If such a discrepancy was perceived, the students in such schools might see the school as inimical and frustrating. Students would tend to use the instrument of redress which was safest and most readily available to them — mob action or group protest — to demonstrate their dissatisfaction and to effect a change. The degree of perceptual congruence in the way the principal and the students saw the school administration was the major independent variable of this investigation. Pepitone (1970) had suggested some implications which this investigation would have by stating:

If such a close relation exists between perception and action, inappropriate or maladaptive social behaviour could be supposed to depend upon incorrect or distorted perceptions of the social situations. From this point of view many problems of interpersonal relations may turn out to be in some measure the consequences of perceptual distortion. (p. 71)

Another theoretical problem hinges on the nature of the principal-student relationship. The relationship between the principal and the students was



conceived as approximating what Jones and Thibaut (1958) referred to as “asymmetrically contingent interaction” (p. 156). The behaviour of students is mostly contingent on the principal’s behaviour and the behaviour of the principal is (as most principals see it) partially contingent on student behaviour, but generally independently determined. Therefore, the degree of mutuality between the principal and the students in their awareness of the apparent asymmetry in their interaction would influence their perception of the principal’s role. Incompatible perceptions of roles may lead to severe strains in social organizations because while one actor may perceive the relationship as fully contingent, the other(s) may perceive it as noncontingent (Hargreaves, 1975). It was also hypothesized that, in the extent to which the principal and the students agree on their perceptions of the administrative role of the principal, to that extent would there be harmony and lower level of student conflict in the school. These hypotheses are diagrammatically shown on the theoretical model given in Figure 1.

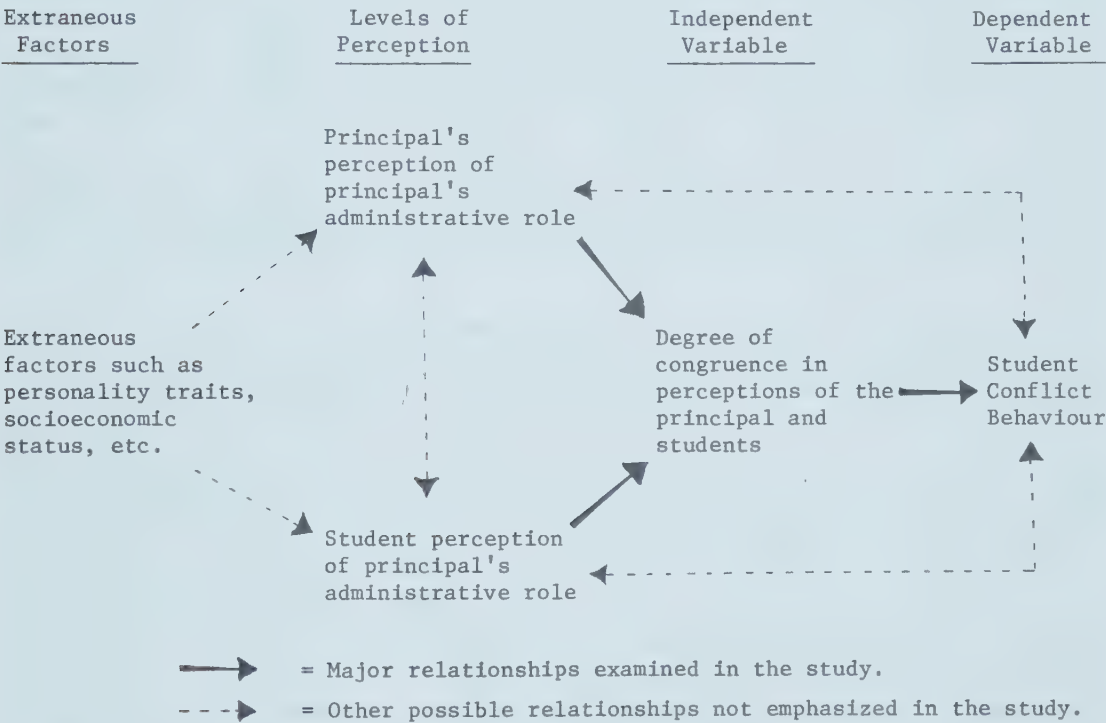


Figure 1. A hypothetical model of the relationship between congruence in principal-student perception of the principal’s administrative role and the level of student conflict.

*The Principal’s Administrative Role.* Generally, the term administrative role would imply the set of norms and expectations applied to incumbents of administrative positions. But when it comes to operationalizing the concept of administrative role for the purpose of measuring it in public and “domesticated” organizations like the school system (Carlson, 1965, p. 3), there arises the problem of which behaviour is purely administrative and not something else. The problem that has plagued the definition of administrative role may be attributable to its historical development. The term administration has been variously and often contradictorily used by various camps (business, commerce, politics, law, and now, education). Secondly, the development of educational administration in Nigeria

under separate controls by different agencies such as religious organizations, the colonial government, and some enterprising local individuals, made it difficult for a clear and precise definition of the principal's administrative role in the school system. The principal's administrative role in Nigeria, therefore, tended to depend on what the principal found himself doing to get the school going or otherwise what the school board for a district had instructed the principals to do. To solve the problem of defining principalship role, the principals were asked to define what they did as school administrators. The Principal Role Perception Questionnaire (PRPQ) Forms I and II were administered to them (Nwankwo, 1977, pp. 289-294). Gross, Mason, and McEachem (1966) had advised that "in studying a particular role, an investigator would presumably try to elicit from the members of a specified population the expectations which they hold for the incumbents of a specified position" (p. 60). From the responses of the principals to the PRPQ Form I instrument, the following administrative role dimensions appeared to be commonly identified by most school principals: (a) planning and policy making; (b) executive action, and (c) control. The principals were also involved in determining the twenty-four items that encompass the above dimensions (PRPQ, Form II), and which comprise the final instrument (PRPQ Form III). Therefore, the definition of the administrative role of the principal in this study is in relation to the principals' behaviours related to planning and policy making, executive action, and control, as these apply to the secondary schools. This definition appeared to agree with Niles' (1941) definition, and Stephen Knezevich First Order Abstractions in the Dual Classification of Terms Describing the Administrative Process (Knezevich, 1975, p. 33).

*Student Conflict.* Although our theoretical model appeared to approximate student conflict behaviour with some measures of "indiscipline", it did not intend to equate the two concepts. Indiscipline tends to connote deviance, delinquency, or other aberrant individual behaviours which cannot adequately reflect the group behavioural contingencies envisaged by the design of this study. Rather, the term "student conflict" is employed because it is apparently a less loaded and empirically more determinable concept than indiscipline. Student conflict refers to actions by groups of students which indicate the existence of incompatible activities, motives, and perceptions. These actions would include activities that obstruct, interfere with, disrupt, or hamper other normal activities in the school. Such conflict behaviours might be indicated by the closure of the school, expulsion, restriction or mass punishment of groups of students, or boycott of school activities. The emphasis in our definition of student conflict was on what Cattell (1948) calls "group syntality", that is, action judged from the point of view of the corporate or consensual behaviour of students as groups rather than in terms of the sum of individual students' misconduct which could be the results of certain physiological, social, economic, or political factors outside the scope of the investigation.

### *Design of the Study*

#### *Sample*

The study sample comprised 50 schools, 50 principals, and 1,000 students. From the Imo State Schools Management Board (SSMB) a list was obtained of 16 secondary schools that had experienced student crises, riots, or violent demonstrations during the period covered by the study (April, 1975 to June, 1976).



Since these schools had the conflict element that was crucial to the study, they were deliberately included in the sample. The remaining 34 schools were selected by "cluster" method, making sure that at least two schools were randomly selected from the list of full secondary schools (up to Class V) in each of the 16 Local Government Areas (1975). For a school to be included in the sample, it was required that its principal have been in the school throughout the period covered by the study. All the schools were tested on the School Conflict Test (see Instruments); twenty schools (including the 16 schools listed in the SSMB Reports for school crises) scored high on the level of student conflict. These twenty schools were categorized as "A" schools, that is, schools with high level of student conflict. The other 30 schools were categorized as "B", that is, schools with low level of student conflict.

The selection of the principal sample was automatic with the selection of his school. For the student sample, 20 students were randomly selected (10 students from each of Forms IV and V). Eventually the ratio of the sample to population was 39 percent for the schools and principals, and 30 percent for the students (that is, students in Forms IV and V in the State under study). These ratios were considered representative for a survey of this nature.

#### *Study Instruments*

Two main questionnaire instruments were used: The Principal Role Perception Questionnaire (PRPQ, Forms I, II, and III) and the Student Conflict Test (SCT).<sup>1</sup>

##### *Principal Role Perception Questionnaire*

(a) PRPQ Form I was the scale by which principals rated, among enumerated administrative role dimensions, the role dimensions related to their daily duties in the schools.

(b) PRPQ Form II was a check list by which principals checked from a list of administrative duties related to identified dimensions of planning and policy-making, executive action, and control, those duties that characterized their daily behaviours in their schools.

(c) PRPQ Form III was the main questionnaire-type instrument developed from Forms I and II. This instrument was used to test the perceptions of the principals and the students on a five-point, equal appearing, Likert-type scale on 24 items. The PRPQ Form III had a reliability of .881.

*Student Conflict Test.* This test was developed during a pilot study and was used to weight the level of student conflict in each school. For the test, the principals, teachers, and selected students in the two uppermost grades of the schools were asked to indicate how often the following conflict indicators had occurred in their schools: the school was closed down as a result of student crises; some groups of students were punished, suspended, expelled, etc., as a result of serious misconduct; or the students boycotted classes, meals, or other school activities in protest against the school or the school authorities.

All the instruments were administered in person, ensuring a 100 percent return and offering the investigators the opportunity to observe some schools at work and to interview some principals and/or students.



*Data Analysis*

The scores on the PRPQ Form III test for each school were summed up across the 24 items. The means were converted into standardized scores using arbitrary mean 50 and standard deviation 10. All schools scoring 50 and above were regarded as category “A” or “High Level of Student Conflict Schools” while those schools scoring below 50 were regarded as category “B” or “Low Level of Student Conflict Schools.” For the data on the PRPQ (Form III) the Gross et al. (1966, p. 167) method of role consensus analysis was tried and found to give extreme correlations; also some researchers had criticized the Gross method of overall between-position consensus as tending to give spurious correlations (Cronbach & Furby, 1970; Manning and Dubois, 1962). Therefore the Hammer and Dachler (1975, p. 65) formula was used in this study to derive the index of congruence in the perception of the leader as an individual and his subordinates as a group. Using this method: (a) the mean of the student ratings of their perception of the principal’s administrative role for each of the 24 items was obtained for all 20 students for each principal; (b) the principal’s ratings of his perception for each item were correlated with the mean ratings for his students, using the Pearson Product Moment Correlation; and (c) the indices of correlation of congruence were converted to standardized scores (to agree with the standardized scores on the Student Conflict Test scores) using an arbitrary mean 50, and standard deviation 10.

*Findings*

*Degree of Congruence in the Principal-Student Perception of the Principal’s Administrative Role*

Table 1 shows the degree of congruence in the principal-student perception of the principal’s role according to individual schools. From this table it appears that 64 percent (32 out of 50) of the sample schools had high degree of congruence on the perception of the principal’s administrative role. But when considered according to school category, few (25%) in category “A” (high level of conflict) had a high degree of congruence in the perception of the principal’s role. Table 2 shows the relative percentage congruence of the sample school categories. Most schools (90%) in category “B” (low level of conflict) appeared to indicate a high degree congruence. The test of significance of the differences according to school category and based on the total sample is shown in Table 3. According to the table, category “A” schools indicated significant differences in the test of the principal-student between-and within-group perception of the principal’s role ( $t = 2.543$ ;  $F = 6.468$ ,  $p < .05$ ), whereas no significant differences were reported in category “B” school ( $t = 1.411$ ;  $F = 2.016$ ,  $p < .05$ ). On the total sample basis, there were no significant differences in the group means ( $t = 1.702$ ;  $p < .05$ ), whereas significant differences were found in the within-group variances ( $F = 2.890$ ;  $p < .05$ ).

*The Relationship between the Degree of Congruence in the Principal-Student Perception of the Principal’s Administrative Role and the Level of Student Conflict Behaviour*

The individual school scores on both the degree of congruence in the principal-student perception of the principal’s administrative role and the scores on the level of student conflict are given in Table 4. A cross-break table analyzing the scores on the relationship between congruence in perception and the level of

TABLE 1  
DEGREE OF CONGRUENCE IN THE PRINCIPAL-STUDENT PERCEPTION  
OF THE PRINCIPAL'S ADMINISTRATIVE ROLE

School Number and Category	Congruency Correlation Coefficients r*	Standard (z) Scores**	Significance of Scores	School Number and Category	Congruency Correlation Coefficients r*	Standard (z) Scores**	Significance of Scores
01 (A)	.201	33	N.S.	26 (B)	.480	50	Sig.
02 (A)	.220	35	N.S.	27 (B)	.480	50	Sig.
03 (A)	.261	37	N.S.	28 (B)	.512	52	Sig.
04 (A)	.243	42	N.S.	29 (B)	.501	51	Sig.
05 (A)	.250	36	N.S.	30 (B)	.541	54	Sig.
06 (A)	.302	39	N.S.	31 (B)	.560	55	Sig.
07 (A)	.242	36	N.S.	32 (B)	.583	56	Sig.
08 (A)	.311	40	N.S.	33 (B)	.591	57	Sig.
09 (A)	.272	38	N.S.	34 (B)	.582	56	Sig.
10 (A)	.340	42	N.S.	35 (B)	.602	57	Sig.
11 (A)	.366	43	N.S.	36 (B)	.601	57	Sig.
12 (A)	.301	39	N.S.	37 (B)	.611	58	Sig.
13 (A)	.391	45	N.S.	38 (B)	.602	57	Sig.
14 (A)	.280	38	N.S.	39 (B)	.632	59	Sig.
15 (A)	.351	42	N.S.	40 (B)	.651	60	Sig.
16 (A)	.491	50	N.S.	41 (B)	.652	60	Sig.
17 (A)	.503	51	Sig.	42 (B)	.650	59	Sig.
18 (A)	.491	50	Sig.	43 (B)	.701	63	Sig.
19 (A)	.531	53	Sig.	44 (B)	.620	68	Sig.
20 (A)	.503	51	Sig.	45 (B)	.683	62	Sig.
21 (B)	.222	35	N.S.	46 (B)	.663	61	Sig.
22 (B)	.241	36	N.S.	47 (B)	.724	64	Sig.
23 (B)	.261	37	N.S.	48 (B)	.763	66	Sig.
24 (B)	.480	50	Sig.	49 (B)	.802	69	Sig.
25 (B)	.491	51	Sig.	50 (B)	.743	65	Sig.

\* N = 24, d.f. = 22; r < .404, p < .05 = N.S.; r > .404, p < .05, r = Sig.

\*\* Standard scores based on arbitrary  $\bar{x}$  50, SD, 10; z < 50 = Low = NS; z > 50 = High = sig. (Ref., Garrett, H. E., *Statistics in Education and Psychology*, 6th Ed., London, 1970, pp. 312-314.)

Total Sample Result: High degree of congruence schools (r > .404; z ≥ 50) = 32 or 64%.  
Low degree of congruence schools (r < .404; z < 50) = 18 or 36%.

TABLE 2  
DEGREE OF CONGRUENCE IN THE PRINCIPAL-STUDENT PERCEPTION  
OF THE PRINCIPAL'S ADMINISTRATIVE ROLE  
ACCORDING TO SCHOOL CATEGORY

Degree of Congruence in Principal-Student Perception of Principal's Role				
	Low		High	
	(r < .404;    z score < 50)		(r > .404;    z score > 50)	
<hr/>				
<u>School Category A (N = 20)</u> (High Level of Conflict)				
Number	15		5	
% Congruence	75%		25%	
<u>School Category B (N = 30)</u> (Low Level of Conflict)				
Number	3		27	
% Congruence	10%		90%	

For r: N = 24, d.f. = 22, p < .05  
For standard score: N = 50

TABLE 3  
TESTS OF SIGNIFICANCE OF DIFFERENCES IN THE PRINCIPAL-STUDENT  
PERCEPTION OF THE PRINCIPAL'S ADMINISTRATIVE ROLE  
WITH RESPECT TO SCHOOL CATEGORY

Sample School Category	MEANS <i>t</i> -test <sup>a</sup>	VARIANCES <i>F</i> -ratio <sup>b</sup>
Category A ( <i>N</i> = 20) (High Level of Student Conflict)	2.543*	6.468*
Category B ( <i>N</i> = 30) (Low Level of Student Conflict)	1.411	2.016
ALL SAMPLE ( <i>N</i> = 50)	1.702	2.890*

<sup>a</sup> d.f.: Category A = 38; Category B = 58; Category A and B = 96.

<sup>b</sup> d.f.: Category A = 1, 38; Category B = 1, 58; Category A, B = 3, 96.

\* Significant differences, *p* < .05

TABLE 4  
INDIVIDUAL SCHOOL RESULTS ON THE DEGREE OF CONGRUENCE ON  
PRINCIPAL ROLE PERCEPTION AND LEVEL OF STUDENT CONFLICT

School Number	Congruency Coefficient <i>r</i> *	Standard Scores on Perception** ( <i>z</i> )	Standard Scores on Level of Conflict** ( <i>z</i> )	School Number	Congruency Coefficient <i>r</i> *	Standard Scores on Perception** ( <i>z</i> )	Standard Scores on Level of Conflict** ( <i>z</i> )
School Category A				School Category B			
01	.201	33	70	21	.222	35	50
02	.220	35	67	22	.241	36	50
03	.261	37	67	23	.261	37	47
04	.243	42	64	24	.480	50	47
05	.250	36	64	25	.491	51	47
06	.302	39	64	26	.480	50	47
07	.242	36	64	27	.480	50	45
08	.311	40	65	28	.512	52	45
09	.272	38	61	29	.501	51	45
10	.340	42	61	30	.541	54	45
11	.366	43	61	31	.560	55	45
12	.301	39	61	32	.583	56	42
13	.391	45	61	33	.591	57	42
14	.280	38	61	34	.582	56	42
15	.351	42	59	35	.602	57	42
16	.491	50	59	36	.601	57	42
17	.503	51	59	37	.611	58	42
18	.491	50	96	38	.602	57	42
19	.531	53	56	39	.632	59	39
20	.503	51	56	40	.651	60	39
				41	.652	60	39
				42	.650	59	39
				43	.701	63	39
				44	.620	68	39
				45	.683	62	39
				46	.663	61	38
				47	.723	64	38
				48	.763	66	38
				49	.802	69	36
				50	.743	65	38

\* All *r*'s are on the Product Moment Formula.

\*\* Standard (*z*) Scores are on arbitrary  $\bar{x}$ , 50, SD. 10. All scores above 50 are significant.



TABLE 5

CROSS-BREAK TABLE OF THE RELATION BETWEEN CONGRUENCE IN THE PRINCIPAL-STUDENT PERCEPTION OF THE PRINCIPAL'S ADMINISTRATIVE ROLE AND THE LEVEL OF STUDENT CONFLICT

Level of Student Conflict	Degree of Congruence in Role Perception		All
	Low (z score < 50)	High (z score > 50)	
Low (School Category "B") (standardized score < 50)	3 A	27 B	30 A+B
High (School Category "A") (standardized score > 50)	15 C	5 D	20 C+D
All Sample Schools ("A", "B")	18 A+C	32 B+D	50 A+B+C+D

$$\chi^2 = 20.00; \text{ d.f.} = 1; p < .05 \quad (\text{See Note 2})$$

$$r_t = -.871, \text{ BC} < \text{AD}, (r_t \text{ is negative}) \quad (\text{See Note 2}).$$

student conflict is shown in Table 5. The results show that there was a strong relationship between the principal-student perception of the principal's administrative role and the level of student conflict in the schools ( $\chi^2 = 20.00$ , d.f. = 1,  $p < .05$ ). Table 5 also indicates that the relationship between the two variables, perceptual congruence and level of conflict, was negative ( $r_t = -.871$ ;  $p < .05$ ), implying that the lower the degree of congruence in the principal-student perception of the principal's role, the higher the level of student conflict behaviour in the schools and vice versa.

### Discussion

Although this study showed interesting results on the degree of perceptual congruence and on the relationship between the perception of the principal's administrative role and the level of student conflict, conclusions were drawn with caution because inferences on organizational behaviour or group social interaction based on perception may be fraught with risks. For group perception of persons or events to be reliable, it has to provide for other dyadic characteristics such as "mutuality" and accuracy. However, by analyzing the scores on the PRPQ Form III Instrument, we got some measures of the degree of congruence in the principal-student perception of the principal's role, and since there were no strong grounds to contest what or how they perceived, we assumed "mutuality" and congruence in their perception. In addition, if mutuality and congruence in perception are assumed, the third characteristic, "accuracy," is necessarily present (Taguiri, 1958, p. 325). The finding in this study of an appreciable degree of congruence in the perceptions of some of the principals and their students appears interesting because most earlier studies had suggested that consensus in perception among groups (even between reference groups) might be difficult to achieve (Gross

et al., 1966, p. 74; Halpin & Winer, 1957; Horowitz, Anderson & Richardson, 1969; Osibov, 1961). The findings of this study may be a consequence of concentration on group perception of stimulus events such as the day-to-day events in the school, rather than on indirect “episodic” or “dispositional” judgements which could be affected by individual personality and social or other extraneous factors outside a rigidly defined school setting.

On the relationship between congruence in the perception of the principal’s administrative role and the level of student conflict behaviour (the main interest of the study), findings showed that most of the schools (75%) indicating a high level of student conflict also tended to have a low degree of congruence in the perceptions of the principal’s role, while a majority of the schools (90%) that indicated a low level of student conflict had a high degree of perceptual congruence on the administrative role of the principal. Several studies in social psychology, sociology, and organizational behaviour have tended to stress the importance of perception in social action. The present study tended to extend that hypothesis to group action by asserting that the nature of group action may also depend on the degree of perceptual congruence on a focal person, event, or situation. People will behave according to their perception of a situation, and apparently, the more consensual the perceptions of people in a social situation, the more functional and harmonious their behaviours tend to be. Our investigation has supported this hypothesis by demonstrating that where there is a high degree of agreement in the way the principal and the students perceive the school administration, there tends to be less student conflict behaviour.

The question may arise as to which of the variables — degree of congruence in the perception of the administration or level of student conflict — causes the other. While it would be necessary to consider the issue of causal relationship between the two variables, this investigation was simply a correlational one and the findings imply that the two variables would tend to be found in association in secondary schools in Nigeria. This assumption presupposes that it could be possible to understand, predict, or control any of the two variables to a reasonable degree by understanding, predicting, or controlling the other. However, it might be easier administratively to understand or manipulate issues related to perception of the school administration in order to influence student behaviour than it would be the other way around.

Another question pertinent to our findings is whether a high degree of congruence in the principal-student perception of the principal’s administrative role comes as an organizational accident or whether it is achieved by administrative design. The answer to this question would draw from our principle of asymmetrical relationship highlighted in the theoretical framework earlier for the study shown earlier. The argument is that student behaviour is directly influenced by the principal’s role, but that the principal’s role (often defined by the School Board) is only partially determined by student behaviour. Logically, therefore, the principal can deliberately influence student behaviour by a carefully programmed design of administration. In other words, the principal has the chance of achieving a high degree of perceptual congruence between himself and his students (and even teachers) by employing appropriate administrative mechanism to help them to perceive his administration as he does. The principal may also achieve a high degree of perceptual congruence through appropriate communication, encouraging student involvement in the school life (House & Miller, 1973, p. 31) and



participation in decision-making. If the members of the school are informed and there is an opportunity to arrive at a measure of consensus on major administrative and other issues affecting the activities and performance of all members, then congruence in perception will tend to be greatly enhanced.

The following characteristics tended to be present in the sample schools where there was a high degree of congruence in the principal-student perception of the principal's administrative role: (a) the staff and students were kept informed of the day-to-day issues, problems, and developments in the school; (b) the principal or his deputy was readily accessible to students much of the time for consultation about their personal problems; (c) the principal occasionally visited students at work and sometimes participated in classroom teaching; (d) there was opportunity for staff and students to participate in decisions related to the day-to-day running of the school; (e) the principal and the teachers showed signs of friendliness, understanding, regard, and confidence; (f) the principal encouraged leadership and innovativeness in both students and teachers; and (g) qualities of leadership, punctuality, and hard work were infused in the school, not by coercion but by the exemplary behaviour of the principal.

In those schools where the principal-student perception of the principal's role was characterized by a low degree of congruence, there were signs of the principal's hoarding information, avoiding group meetings on school matters, being skeptical about delegating authority to staff or students, and tending to overload their office and administrative channels. The result was often a break in communication between the principal's perception of his role and the students' perception of the principal's role, resulting in low perceptual congruence. This situation of communication breakdown between the principal and the students (Erikson, 1966), created room for speculations, rumours, and suspicions, thus setting the stage for school-induced frustration and, finally, student conflict.

### *Conclusion*

Our study suggests that a major determinant of student conflict behaviour in Nigerian secondary schools may be the perception of the school administration. The findings suggest that student conflict behaviour is not necessarily an imperative school phenomenon, but rather a situational consequence of the degree of perceptual congruence in the school. We also suggest that student conflict may be reduced through appropriate administrative mechanisms that enhance perceptual congruence or consensus in expectations, such as improving communication between the students and the school administration.

### *Notes*

1. These instruments were developed during the pilot study on the development of instruments for a Ph.D. research. For the development and sample of these instruments, see Nwankwo (1977), pp. 147-164, and Appendix B<sub>1</sub>, B<sub>2</sub>, and B<sub>3</sub>.
2.  $\chi^2$  (Chi-square) may be obtained from the solution of the  $2 \times 2$  contingency table, and using a table of Chi-square to test the significance: d.f. = 1,  $p < .05$ . The tetrachoric correlation coefficient,  $r_t$ , is more easily derived from the  $2 \times 2$  (or AD/BC) contingency table. This was considered necessary in order to determine both the degree and the direction of the relationship between variables (see Garrett, 1970, pp. 384-388).

### *References*

- Blackstone, T., Gales, K., Hardley, R., & Lewis, W. *Students in conflict: L.S.E. 1967*. (L.S.E. Research Monographs 5). London: Weidenfeld and Nicolson, 1967.
- Carlson, R. O. Barriers to change in public schools. In R. O. Carlson et al., *Change process in*



- public schools. Eugene, Oregon: Centre for Advanced Study of Educational Administration, 1965.
- Cattell, R. B. Concepts and methods of measurement of group syntality. *Psychological Review*, 1948, 55, 48-63.
- Chesler, M., & Franklin, J. Interracial and interorganizational conflict in secondary schools. Paper presented at the Annual Meeting of American Sociological Association, Boston, Mass., August, 1968.
- Coleman, J. S. *The adolescent society*. New York: Free Press, 1961.
- Cronbach, L. J., & Furby, L. L. How should we measure "change" — Or should we? *Psychological Bulletin*, 1970, 74, 68-80.
- Daily Star*. Enugu, Nigeria.
- Daily Times*. Lagos, Nigeria.
- Davis, K. Adolescence and social structure. *The Annals of the American Academy of Social and Political Sciences*, November, 1944.
- Deutsch, M. Conflicts: Productive and destructive. *Journal of Social Issues*, 1969, 25, 7-43.
- Elkin, F., & Westley, W. A. The myth of adolescent culture. *American Sociological Review*, 1955, 20, 680-684.
- Erikson, D. A. (Ed.) *Educational organization and administration*. Berkeley, California: American Educational Research Association, 1966.
- Frankenstein, C. *Varieties of juvenile delinquency*. London: Gordon and Breach Science Publishers, 1970.
- Garrett, H. E. *Statistics in psychology and education*, 6th edition. London: Longmans, 1970.
- Green, A. *Sociology*. New York: McGraw Hill Book Company, 1952.
- Gross, N., Mason, W. B., & McEachem, A. W. *Explorations in role analysis*. New York: John Wiley and Sons, Inc. 1966.
- Halpin, A. W., & Winer, B. J. A factual study of the leader descriptions. In R. M. Stogdill & A. E. Coons (Eds.), *Leadership behaviour: Its description and measurement*. Columbus: Bureau of Business Research, The Ohio State University, 1957, 39-51.
- Hammer, T. H., & Dachler, H. P. A test of some assumptions underlying the Parth-Goal model of supervision: Some suggested conceptual modifications. *Organizational Behavior and Human Performance*, 1975, 14(1), 60-75.
- Hargreaves, D. H. *Interpersonal perception in education* (Revised student edition). London: Routledge and Kegan Paul, 1975.
- Heider, F. Consciousness, the perceptual world, and communication with others. In R. Taguiri & L. Petrullo (Eds.), *Persons perception and interpersonal behaviour*. California: Stanford University Press, 1958.
- Hollingshead, A. B. *Elmonts youth*. New York: John Wiley, 1949.
- Horowitz, M., Anderson, G. J., & Richardson, D. N. Divergent views of the principal's role: Expectations held by principals, teachers and superintendents. *Alberta Journal of Educational Research*, 1969, 15(4), 195-206.
- House, J., & Miller, W. Responding to student unrest. *Curriculum Bulletin*, 1973, 27, January, Oregon ASCD No. 315.
- Jones, E. E., & Thibaut, J. W. Interaction goals as bases of inference in interpersonal perception. In R. Taguiri & L. Petrullo (Eds.), *Persons perception and interpersonal behaviour*. California: Stanford University Press, 1958, 151-177.
- Knezevich, S. J. *Administration of public education* (3rd ed.). New York: Harper and Row Publishers, 1975.
- Linowitz, S. M. (Chairman). Campus tensions: Analysis and recommendations. Report of the Special Committee on Campus Tensions. Washington, American Council on Education, December, 1970.
- Manning, F. W. H., & Dubois, P. H. Correlational methods in research on human learning. *Perceptual and Motor Skills*, 1962, 25, 287-321.
- Miklos, E. The role theory in administration. *Canadian Administration*, 1963, 3, November, 5-8.

- Nigerian Chronicle*. Calabar, Nigeria.
- Niles, M. C. H. *Middle management*. New York: Harper and Row Publishers, 1941.
- Nwankwo, J. I. Perception of the administrative roles of the principal and the school climate as factors in students' conflict behaviour. Unpublished doctoral thesis, University of Ibadan, Ibadan, Nigeria, 1977.
- Oloruntimehin, O. Some factors related to juvenile delinquency in Nigerian children. *West African Journal of Education*, 1974, 17(2), June, 117-122.
- Osibov, H. Professional and public perceptions of superintendent behaviours. Doctoral thesis, University of Oregon, Eugene, Oregon, 1961.
- Pepitone, A. The determinants of distortion in social perception. In H. Proshansky & B. Sidenberg (Eds.), *Basic studies in social psychology*. London: Holt, Rinehart and Winston, 1970.
- Rogers, C. R. *Freedom to learn*. Ohio: Merrill, 1969.
- Taguiri, R. Social preferences and its perception. In R. Taguiri & L. Petrullo (Eds.), *Persons perception and interpersonal behaviour*. California: Stanford University Press, 1958.
- Waller, W. *The sociology of teaching* (Science Ed.). London: John Wiley and Sons, Inc., 1965.
- Williams, R. *The American society*. New York: Alfred Knopf, 1952.
- Zalkind, S. S., & Costello, T. W. Perception: Some recent research and implications for administration. *Administrative Science Quarterly*, 1962, 7, September, 218-235.

M. K. BACCHUS

*The University of Alberta*

## Education as a Social Control Mechanism

*During slavery, West Indian planters were opposed to the education of slaves. They felt that this might make the slaves more discontent with their conditions and hence more difficult to control. Instead the planters chose to rely on coercive methods of control to ensure compliance with the existing order of the society. However, with the proposed abolition of slavery, they completely changed their attitudes towards education for the black population. They then hoped that with the help of the churches, education would succeed in developing among the black population, voluntaristic support for the exploitative features of slave society even after its legal props were removed by emancipation. (Dr. Bacchus is a Professor in the Department of Educational Foundations at The University of Alberta.)*

### *The Role of Education in Slave and Post-Slave West Indian Society*

The existing historical evidence indicates that just before and immediately following the emancipation of slavery, there was a dramatic change in the attitudes of the majority of West Indian planters towards education for the black masses. This paper will (a) examine why this marked attitudinal change occurred; (b) throw some light on how the planters perceived the role of education among the black population; and (c) speculate about the use of education as a mechanism of social control in societies which are increasingly extending citizenship rights to all sections of their population.

However, before discussing these issues, it will be useful to identify some of the essential features of slave society which influenced the attitude of the planters towards the education of the black masses.

### *Some Features of Slave Society*

The major structural features of West Indian society which were mainly responsible for the negative attitude of the planters towards the education of the slaves were:

1. It denied membership in the political system to the slaves who formed the numerically largest population group in these countries.



2. It not only permitted the slaves to be owned like chattels, but sanctioned the whole system of economic production which was based on the exploitation of one group solely in the economic interests of another.

3. Total control of the state machinery rested with the plantocracy. Marx's references to the state as the "executive committee of the ruling class" quite aptly described the political arrangement in slave society.

The major aim of the West Indian planters was to maintain the economic, political and social arrangements of slave society especially since the privileges which they enjoyed from it were phenomenal. The planters were afraid that any change in the system of slavery might start what Myrdal (1964) refers to as a "process of cumulation" which might eventually affect their position of dominance in the society. That is why they continuously opposed most of the changes suggested by the Colonial Office for ameliorating the harsh conditions under which the slaves lived and worked. The editor of the *Colonist* was putting over the views of most planters when he argued that "slavery must exist *as it is now* or it will not exist at all [emphasis added] (*The Colonist*, 1824).

The two major means by which the planters attempted to ensure the stability of slave society were: (a) the use of coercive power over the slaves, both in their role as slave owners and as managers of the state-control mechanisms; and (b) the use of various means of legitimation to bring about a great degree of acceptance of, or at least acquiescence to, the existing political, social and economic system, by the slaves, e.g., through socialization practices directed both at locally born and newly imported slaves, and enacting supporting laws to ensure almost complete dependence of the slaves on the plantations.

A third method that could have been used to help develop legitimacy for the society was possibly a fairer distribution of its wealth, but this would have necessitated changes in the nature of the social order, changes which the planters were understandably unwilling to accept.

While planters relied heavily on the first mentioned mechanism of social control, i.e. the use of coercive power, they were generally aware of the importance of the second mechanism — "proper" socialization. From their own experience they recognized the inherent instability of a social order built upon the sheer use of coercive power and tried to socialize slaves to accept their positions of bondage and the dominance and superiority of their masters. Scholars such as Patterson (1967) have described the socialization process which occurred among creole slaves and the methods used for "seasoning" new slaves recently arrived from Africa. The objective was to develop in them not only the lower-level job skills which were required on the sugar plantations, but also the desired attitudes towards work and to the social order.

Since schooling is often considered one of the more effective institutionalized means of socialization and social reproduction, i.e., of passing on to the young generation those values, beliefs, and social relationships which are supportive of the existing order of society, the question is "Why did the planters not turn to education in this situation? Or more correctly, "Why were the planters so strongly opposed to the education of slave children?", despite the possibilities inherent in schooling of inculcating attitudes and beliefs which might have helped to bring about social stability.

There were many reasons for this. First, education performs both a conservative

and, if not a radical, at least a liberating function. In addition to passing on those values, beliefs, attitudes, and skills which are needed to give support to the existing social order, it can also open up new vistas or new horizons among its recipients. To the planters this meant that if education was made available to the slaves, it might have led some of them to question even more vehemently the morality of a society in which gross inequalities based on race were so firmly institutionalized. As major beneficiaries of the existing system, planters were perceptive enough to know that it was dangerous to allow slaves to acquire knowledge and ideas which might lead them to chafe more heavily under their position of bondage. Rev. John Smith (March 18th, 1818) had observed that generally the planters felt that "to teach the slaves is an impolitick measure" while Gordon (1963, p. 3) noted that one of the planters' objections to slaves learning to read was the realization that "reading provokes thinking and this is dangerous in a slave society." An important concern of the planters was to prevent the aspiration of the slaves from rising, or their feeling of deprivation from increasing. This is why they saw education and slavery as being incompatible.

In addition, the governing classes in Britain who formed the major reference groups for the local planters, were still not convinced about the role of education in pacifying the masses. In fact, it was only in 1833 that the House of Commons approved its first meagre grant to aid the efforts of the voluntary agencies in providing education for the children of the British working class.

Another reason why slave-owners objected to the education of slaves was the fear that once they had acquired literacy skills, it would become even more difficult to monitor their access to new information and knowledge. More specifically, they were afraid that if the slaves learned to read, they might find out about the efforts of the abolitionists in England to end slavery and it was realized that such information might increase their discontent and rebelliousness. Even a more liberal plantation owner like Hermanus Post, who brought the Rev. John Wray to minister to the spiritual needs of his slaves, made it quite clear to the missionary that he should not teach the slaves to read. Similarly, when Wray's successor, the Rev. John Smith arrived in Demerara in 1817, the Governor warned him that if he taught a slave to read and he heard of it, he (Smith) would be banished from the Colony (Daly, 1966, p. 135). Because of this fear of slaves learning to read, the missionaries were expected to use only oral means of instruction in teaching them about religion.

A third reason for the planters' objection to the education of slaves was the fact that most of them had a great distrust for the work of the missionaries, especially the evangelical missionaries, and felt that if a check was not put on their activities they would prove "subversive to good order." They realized the dangers of teaching the slaves that they were equal to the whites in the sight of God, because this might eventually lead them to question not only the gross inequality in the society, but also the moral right of one individual to own and exploit other individuals who were spiritually their equals. According to the *Colonist* (Feb. 18th, 1824), "it is not a matter of surprise that a Negro slave who is taught that all men are equal from a religious point of view should wish the same principle to prevail in politics."

The planters regarded the efforts of the missionaries as being directly responsible for increasing social unrest in the society and, more specifically, for a number of slave rebellions that occurred. The most notable incident was the accusation and trial of Rev. John Smith for complicity in the 1823 East Coast Slave



Insurrection in Demerara. In view of this, the planters felt that any attempt to institutionalize the work of these missionaries through the establishment of schools might sound the death knell of slave society.

The editor of the *Colonist* expressed the views of the planters very clearly when, following the death of Rev. John Smith, he wrote an editorial castigating planters for not speaking out in time against the first advocates of missions and education for the slaves. He pointed out that these advocates should have been told that “the missionary system is, in fact, undermining the institutions and endangering the political existence of the colonies.” Continuing, he (Daly, 1966) noted:

While we have no desire to treat Africans with undue rigour . . . we cannot be ignorant (of the fact) that our power over them can exist only so long as we are more *highly educated* and *enlightened*: We are few; they are many; and if their moral qualities or *education be allowed to equal ours*, it follows that the *power* or the *right* of government *which is the same thing*, will be determined by the amount of physical force [emphasis added]. (p. 254)

This fear that the education of slaves would lead to social unrest and “disorder” was shared by slave owners in other societies. In the United States of America, laws were passed which expressly forbade the education of slaves. Also, in the French West Indies, the Governor of Martinique was using the same argument as the British West Indian planters in trying to prevent the extension of religious instruction and education to the slaves there. In semi-private correspondence to his department in France, he wrote:

I came here with all the European prejudices in favour of the necessity of instructing Negroes in the principles of religion. But sound policy and even more powerful consideration of humanity are opposed to this. The *safety of the whites demand that they should be kept in the most profound ignorance*. . . . The *safety* of the whites, fewer in number, demands ignorance on the part of *the slaves* [emphasis added]. (Augier & Gordon, 1962, pp. 145-146)

H. Coleridge (1832), who made a six-month visit to the West Indies in 1825, summed up thus the planters’ dilemma in the matter of the education of slaves:

If the planters attempt to educate the slave, they do too little for their own safety in persisting to debar him from the privileges to which he would soon feel that he has required an equitable right. *It will be impossible to march Negroes on the road to knowledge and compel them to stand at ease within the old entrenchments of ignorance* [emphasis added]. (p. 321)

Coleridge was very perceptive about the planters’ concern that education of slaves was likely to add another disequilibrating element in the society since it would raise their levels of aspiration which, if not met in some way, would have had serious disruptive effects.

The planters also had an immediate economic interest in not wanting slave children to attend school. At the age of six, these children were already working in the grass gangs making an economic contribution which was enough to cover expenses involved in their upkeep. If they were to be sent to school, the planters had to bear that part of the cost of their education which would have resulted from their lost contribution to production. Also, it was necessary to inculcate good work habits in these young slaves and it was thought that sending them to school might have the opposite effects. As Shirley Gordon (1963) noted:

Keeping slaves busy was one of the most reliable means of keeping them acquiescent. The



principle was applied early in the life of a slave born child; schooling would have been a contradiction of the whole system. (p. 10)

Finally, part of the planters' objections to the education of the slaves might have resulted from their general attitude to the black and coloured population, whom some of them saw as "barbarians," a "good for nothing race of beings" and who, according to the Governor of Martinique, ought to be treated "as one treats beasts." But against this attitude, there was evidence of the ability of the black population to do well educationally since many of them had already qualified as professionals, especially as lawyers, after studying in British institutions of higher education. So though there were some doubts about the educability of the black population, these were not as widespread as in other societies.

Incidentally the planters were not opposed to allowing slaves to acquire training in specific skills needed on the plantations. Despite the restrictions imposed by some West Indian Governments against allowing slaves to do skilled work, most of the estate craftsmen such as boilermen, coopers, carpenters, and masons were slaves and obtained their training for these jobs within the slave system. Because of the superior status and additional privileges enjoyed by these tradesmen, many slaves avidly sought the opportunity, either for themselves or for their children, to be apprenticed to a tradesman. In addition, they were prepared to make the necessary "investment" to obtain such training by giving up their spare time and the opportunity they had to work on their own plots of land in order to cultivate the farms of those slaves to whom they were apprenticed.

While the planters facilitated this manpower training function of education, they were opposed to giving slaves access to the content of knowledge and information, and to their acquisition of literacy skills which would have opened up to them possibilities for further learning. They felt that the likely consequences of this would be detrimental to the stability of slave society. It might give slaves a new dimension to their thinking, raise their levels of social, economic, and political aspirations and lead them eventually to question and challenge even more actively the legitimacy of the society in which they were considered chattels existing mainly for the economic benefit of their masters. The inevitable sequence, as the planters saw it, was "read the bible, then the newspaper till they grow discontented, rebel, and burn the country" (Waddell, 1863, p. 195).

In summary, it was not that the West Indian planters were entirely unaware of the possible role of education in helping to produce a submissive black population and a pliant work-force, but that they were very hesitant to take what seemed to them to be the great risks that might have accompanied the formal education of slaves. As Rooke (1977) indicated

The planters were as unreceptive to an education which taught any form of equality between slave and master as one which included reading. They saw social stability being threatened as much by the assumptions behind missionary education as by the practice of it. Doctrines of spiritual equality exposed the contradiction inherent in slave societies. As the planters correctly recognized, teach a slave he is lovable in the sight of God and he was bound to become "uppity" . . . Teach a slave that Christ came to save all men and he will begin to wonder why some men deserve to own slaves while others deserve to be slaves. (p. 196)

Rev. John Smith came to a somewhat similar conclusion when he reported to the Directors of the London Missionary Society in 1818 that even those planters

who did not question the original intention of the missionaries to make better slaves and inculcate submission and similar dispositions feared that education would “ultimately have a bad effect. The most innocuous education was merely a first step to creating a more self-conscious group of subservients.” In summary, the planters knew that “Christianity and slavery cannot long exist together” and were therefore “resolutely opposed to any form of instruction” for slaves (Rooke, 1977, p. 197).

There were also problems which adversely affected the degree of success of the planters’ efforts at producing a docile slave population. First, those slaves who were newly arrived from Africa were not socialized in their early days to accept a life of bondage and many of them actively showed their resentment to being kept as slaves. Second, as C. L. R. James indicated, it was difficult to subdue the slave’s dream of freedom. During their midnight celebrations, they would often sing their favourite song which, translated into English meant, “we swear to destroy the whites and all that they possess; let us die rather than fail to keep this vow” (James, 1973, p. 16).

Third, there was never, and in the circumstances unlikely to be, any consensus between the two groups to the *raison d’être* of the society — not even any shared view of the complementarity of their separate ends. The social order was a good example of what Hobbes saw as the ultimate outcome of societies in which one group controlled all the power without any major institutional restraints on their desires or greed. Furthermore, there was no shared belief system or ideology which would have justified or rationalized the existing distribution of wealth and power.

Fourth, the difficulty of getting the slaves to accept the existing economic order was aggravated by the contrasting conditions of life between the two groups — the vast wealth conspicuously consumed by the planters as compared to the conditions of degradation experienced by the slaves. The feeling of deprivation must have been further enhanced by the fact that the slaves saw whites who formerly worked alongside them, mainly as indentured labourers, rising up the economic and social ladder once they had obtained their freedom. Commenting on circumstances like these, Moore (1963) pointed out:

On the evidence, it appears extremely unlikely (a cautious way of saying “impossible”) that any mode of rewarding positions unequally and any mode of determining access to those positions should be so firmly institutionalized that those persons most injured would accept the justice of their fate . . . the rules governing assignments to positions and their unequal rewards, and the values that “justify” these rules, will not be accepted as totally valid by those who are thereby excluded. (p. 83)

With the understandable absence of any shared normative order to regulate the inevitable “war of all against all,” and the relative failure of efforts at socializing the slaves to accept their position of bondage and their exclusion from citizenship, the planters depended heavily on the raw use of power in their attempts to maintain order in the society. As James (1963, p. 4) pointed out, it became necessary to create “a regime of calculated brutality and terrorism” in order “to cow them (the slaves) into the necessary docility and acceptance of the social order.” The concentration of political power in the hands of the planters gave them almost monopolistic control and legalized use of the means of violence at the command of the state, and historians report that these were often ruthlessly used to suppress any incipient signs of group or even individual aberrant behaviour.



In short, with the perceived ineffectiveness or failure of the methods of socializing the slaves to accept the existing order to society, the planters depended predominantly on physical control mechanisms and the legalized use of violence to prevent social disruption.

### *Emancipation and After*

By the 1830s, it was known that the Government of the United Kingdom was going to abolish slavery throughout the British Empire and while the planters had considerable control over their local legislatures they were powerless to prevent the introduction of such a bill in the British Parliament. This new situation posed a problem for the planters and helped considerably to bring about a change in their attitudes towards education for the mass of the black population. The planters wanted to maintain some of the basic features of slave society and especially the exploitative relationship which existed between themselves and the black masses — even after the legal props which supported the system of exploitation were withdrawn.

It was feared that the act for Abolition of Slavery might result in the slaves taking steps to prevent this further exploitation by the planters, even if this meant removing themselves from the sugar plantations and settling elsewhere in the country. In addition, they might no longer be willing to accept, without challenge, the superior social and economic position which the whites traditionally enjoyed in the society — a factor which gave support to the belief among the nonwhites in the “natural superiority” of the European and his culture.

The planters therefore took a number of steps to ensure the continued dependence of the slaves on the plantations. For example, they refused to sell or lease land to ex-slaves who, during the period of apprenticeship, had saved some of their wages for this purpose. Even if some planters did, they asked exorbitant prices for their land making it difficult for an economically viable peasantry to develop. In addition,

the provision grounds, plantation walks and fruit trees cultivated and reared by the peasantry while in bondage were destroyed on many estates in order to control their (the peasants') dependence on plantation work. (Martin, n.d., p. 175)

Another important step taken by the planters in this direction was to attempt to build up voluntaristic support on the part of the masses for the exploitative system. There were, among others, two specific and important commitments which had to be inculcated. One was to develop or strengthen the idea that hard work as an estate labourer was a social and moral obligation on the part of the labouring classes. The second was a belief in the superiority of the European and his culture and an acceptance of the view that only Europeans, or those most like Europeans in their values, had the right to occupy positions in the upper strata of the society.

In other words, the planters were now concerned with legitimating the existing social and economic order of “plantation society,” essentially slave society without legalized slavery, by inculcating normative support for it on the part of the masses. For this they began to turn to education. In fact there were a number of reasons for this change in attitude to the education of the black population.

First the Act for the Abolition of Slavery, to which they were opposed, was introduced above their heads by the British Government, adding an important new variable into the situation. Since they could no longer depend on the legalized use



of the coercive power of the state to ensure compliance by the masses, planters were forced to look at alternative means of social control. Developing normative support for the social order through education thus became a much more attractive proposition.

Second, the Act of Emancipation included a grant of £30,000 per annum for five years to promote education among the ex-slaves, and the Imperial Government had decided that the "religious and moral" education which "must form the basis and must be made the inseparable attendant" of any education "for the Negro population to be emancipated" was to be provided "through the agency of the different religious bodies already engaged in promoting education in the Colonies" (Gordon, 1963, pp. 20-21). Therefore, not only was the provision of education for the masses a *fait accompli* as far as the planters were concerned but so was the involvement of the churches in these efforts. While the planters were very critical of the educational efforts of missionaries, they had much more faith in the work of the established churches and, in some islands, the Moravians. It was therefore in their interest to ensure that these particular religious groups were given maximum support in the provision of education among the masses.

Third, the increasing willingness by the British ruling class to accept the view that "appropriate" education provided by the more conservative religious denominations could further assist in the domestication of the "lower orders" would also have influenced the planters' attitudes to the education of the black masses.

Education for this group contributed to social stability in another way. It was unlikely that any section of the population would willingly accept its position on the lowest rungs of the social and economic hierarchy especially if, as in the case of the ex-slaves, they were in the majority, unless they saw the possibility of later social advancement for themselves or their children, or unless they shared some religious or other ideological belief which restrained their ambitions. Even though social mobility was difficult and, for the majority of the black masses, well nigh impossible, education held out some promise of social advancement and gave its recipients a hope, unrealistic though it might have been for the great majority, of rising somewhat up the social and economic ladder. While this was not the purpose of education as conceived by the planters, it nevertheless became the most important one to the ex-slaves who valued education mainly as an instrument of upward social mobility for their children.

(1) The first major role of education as seen by the planters was therefore to produce a pliant and submissive work force, a view shared by the more influential Christian churches of the day. This function of education was also strongly supported by Rev. J. Sterling who, during the apprenticeship period, was appointed by the British Government to review the educational situation in the West Indies. In his report, he emphasized the importance of education as an instrument which would help the planters develop such "power over the minds of the labouring classes" that a continued supply of labour for the sugar plantations would be ensured. He noted that:

The peace and prosperity of the Empire at large may not be remotely influenced by their (the ex-slaves) moral condition. . . . Their performance of the functions of the labouring class in a small civilized community will depend entirely on the *power over their minds* of the same prudential and moral motives which govern more or less the mass of the people there. If they are not so disposed as to fulfill these functions (as labourers), property will

perish in the colonies for lack of human impulsion. . . . The certain result of the new situation when the minds of people are all in movement will be a *consciousness of their own independent value as rational beings without reference to the purposes for which they may be profitable to others*. The law having determined and enforced their civic rights the task of bettering their conditions can be further advanced only by education [emphasis added]. (Gordon, 1963, pp. 20-21)

As can be gleaned from the above quotation, there were then three major concerns in providing education for the masses.

(a) It was to assist in producing labourers who would be docile, obedient and willing to continue working on the sugar estate. This view was also advanced in the *Colonist* — a newspaper which had previously objected strongly to the education of the slaves. Its new argument was that education:

Should be directed to motivating the Africans to work and respect authority. . . . They should learn that idleness is a crime against a higher law as well as an infraction of the social duty and that the commandment which tells them to keep holy the seventh day is equally explicit in requiring them to labour during the six. (Daly, 1966, p. 264)

This attitude was shared by the dominant groups in the other West Indian territories. For example, in January of 1892, about sixty years after the abolition of slavery, the *Agricultural Reporter* was still bemoaning the fact that:

In our elementary schools, we are teaching Geography and Grammar and History and Singing and so forth. . . . We are neglecting studies which would teach the rising generation, not only the dignity of manual labour but how to *labor freely to get their living in that sphere of life into which it has pleased God to call them*. . . . The sons and daughters of the labouring classes have been impregnated with the idea that the education which is imparted to them makes them better than their fathers; and they have begun to look with disdain upon mere manual labour. . . . Unless steps are taken to put an end to this state, the most disastrous consequences must ensue [emphasis added]. (Gordon, 1963, p. 135)

Even Her Majesty's Government which provided money for the education of the children of ex-slaves, had hoped that "the labouring classes will be animated by the same spirit of steady and patient industry which ought always to accompany good instruction" (Circular Dispatch, 1st Oct. 1838). Latrobe, the Inspector of Schools sent out by the Imperial Government in 1837 to report on the results of the first two years of the Negro Education Act used the same criteria in assessing ongoing educational programs. "I shall," he remarked,

be inclined to doubt the wisdom or real kindness of any system or mode of instruction which will lead either the parents or the child to reason falsely on the subject *by not strongly impressing upon their minds the necessity* of submitting to labour [emphasis added]. (Gordon, 1963, p. 30)

(b) A second concern was to teach the masses to accept their place in society and not attempt to upset the status-quo. Again the *Colonist* argued,

Education should break down false ideas of independence which teach the Negroes to disregard the diversities of rank and condition of life imposed for wise purposes. They should be brought practically to love honesty, sobriety, reverence to authority and a Christian respect for all whom Providence has placed in a superior condition. (Daly, 1966, p. 264)

The first Inspector of Schools for British Guiana knew that the ruling class perceived elementary education for the masses as an instrument of social control



and capitalized on this fact whenever he attempted to secure more funds for education. He argued that “money laid out for the encouragement of virtue and the removal of ignorance is a far more profitable investment than that spent on the repression of punishment and crime” (Special Report, Government of Great Britain, 1901, p. 754).

And in the same vein, the editor of the *Royal Gazette*, another local newspaper which represented the views of the planters, reasoned that since the Government had accepted the responsibility for institutions such as prisons, the police force, and the courts, “it would be justified in building schools and paying teachers to prevent crime” (Daly, 1966, p. 264). Since almost any transgression against the planter class was, at the time, likely to be considered a crime, it was obvious that the editor was concerned with the potential role of schools in preventing the spread of deviant attitudes and behaviours among the black masses.

(c) A third concern was the “religious and moral conditions” of the ex-slaves. This could, for example, be seen in the efforts of the Colonial Secretary to get the West Indian assemblies to introduce compulsory education for the masses, at least three and a half decades before such legislation appeared in England. In his correspondence to the West Indian Governments, the Secretary of State for the Colonies (1835) argued:

Whatever objection may exist in more advanced societies to the principle of compulsory education they can have no place in reference to a colony in which the great mass of the people have just emerged from slavery but have not yet generally acquired any acquaintance with the principles and precepts of Christianity and are for the most part destitute of the first elements of learning. . . . The advantages (of education) are so great (for them) that if necessary, the members ought to have it forced upon them by legislative process.

(2) The second major role of education as perceived by the ruling groups was to develop or strengthen among the ex-slaves a belief in the natural superiority of the European and his culture. This was important to give legitimacy to their position as rulers of the society occupying the topmost rungs of the social and economic hierarchy. Overall, West Indian schools were very successful in this role and helped to produce a West Indian elite who not only accepted the superiority of English culture but in most ways became black Englishmen. Smith (1967) observed that one of the basic facts of West Indian creole society was that it was integrated around the conception of the moral and cultural superiority of things British.

These efforts at socializing the black West Indian masses to accept the cultural superiority of the Europeans were also manifest in French West Indian islands. For example, Fanon (1967), who was himself from the French West Indian island of Martinique, observed:

Until 1936, the West Indian lived, thought, dreamed, composed poems, wrote novels exactly as a white man would have done. . . . The West Indian identified himself with the white man, adopted a white man’s attitude, was a white man. (p. 26)

It was these new roles which education was expected to perform in “civilizing” and “tempering the masses” which contributed substantially to the change in the attitudes of West Indian planters towards the provision of education for the children of the ex-slaves. Therefore, just before and immediately following the abolition of slavery, educational activities in the British West Indies were intensified, often with the active support, or at least with little or no opposition



from the planters who, a few years previously, had strongly objected to the education of the children of slaves.

### *Summary and Conclusions*

During slavery the sugar planters strongly opposed formal education for the slaves. They were not interested in its possible contribution in socializing the masses to accept, or at least to be reconciled to, the values which underpinned that society. The potential dangers of educating slaves were too great for the planters to consider it a feasible course of action. They believed that it might have opened up new horizons to the slaves, raised their level of social and occupational aspiration, made them “uppity”, and increased their dissatisfaction with the conditions under which they worked and lived. It was further thought that education might have led the slaves to challenge even more actively “the right of government” to uphold the institution of slavery which allowed the planters to own and exploit them entirely in their own economic interest. As the editor of the *Colonist* argued, “our (the planters’) power over them (the slaves) can exist only so long as we are more highly educated and enlightened” (Daly, 1966, p. 254).

Instead, the planters depended partly on their own methods of socializing or “seasoning” the slaves and more heavily on the use of physical violence to ensure conformity and social stability. In this they had the full support of the law which allowed them almost indiscriminate use of physical punishment to cow the slaves into submission and subservience. Planters did, however, permit slaves to acquire skills needed for their work on the plantations through an informal system of apprenticeship.

But with the proposed abolition of slavery there developed the fear that unless something was done to get the slaves to accept “voluntarily” the major features of plantation society they would, as Rev. Sterling (1835) had warned, “become conscious of their own independent value as rational beings *without reference to the purpose for which they may be profitable to others.*” The ultimate result of this would have been that “property will perish in the colonies . . . [and] the whites would no longer live there.” To prevent this “disastrous” situation from developing it was argued that the “*systematic provision (of educational facilities) for their mental improvement*” and for developing a “*wholesome attitude to work*” was needed. [emphasis added].

It was in the face of this dilemma that the planters began supporting education for the black masses. In this they received the active support of most of the Christian denominations, especially the established churches. The churches and the schools reinforced each other in their attempts to socialize the ex-slaves to perform well “the functions of a labouring class in a civilized community” (Sterling Report, 1835). As R. T. Smith (1962) noted, “schools and churches were held to be the best instruments for the transformation of a rebellious slave population into a peaceful and obedient working class” (p. 145).

The schools were largely successful in these efforts and helped to produce a West Indian population who, as Fanon (1967) pointed out, were European in their values, behaviours, and even in their ways of thinking and expressing themselves. There were, however, important unintended consequences of these educational efforts which also produced “deviants” who were later influential in bringing about political changes in these societies which ultimately led to the transfer of formal

political power from the white minorities to the elected representatives of the masses.

### *Some Speculative Observations*

In any society, the maintenance of social stability depends jointly on: (a) how successful the instruments of socialization are in getting nearly all sections of the population to accept the legitimacy of the existing social order and its institutionalized ways of dealing with opposition; and (b) how effective the control mechanisms of the state are — the police, the armed forces, the courts and sometimes even the news media — in containing disruptive elements in the society. These state control mechanisms essentially attempt to ensure that those “deviants” who have not been “properly socialized” to accept the existing social order or to use the institutionalized means of bringing about changes are rendered ineffective in their efforts to upset the existing society by “illegal” means.

These two methods of social control have to be seen as complementary — the more “effective” the process of socialization the less need there is for the active use of the other control mechanisms of the state.

Education, because of its dominantly conservative role, has generally been considered one of the more useful means of socialization and many educational philosophers have argued the point that the more efficient the school system is, the less need there will be for other institutions of state control, such as the army, to ensure internal social stability. For example, as Massialas (1969) noted, “the German philosopher and educator Fichte made a strong case for the nation to invest in education because education would fulfill the same purposes that standing armies fulfill; viz the support and maintenance of the State” (p. 2).

The relative emphasis which a nation places on education against other social control mechanisms will no doubt be influenced by (a) the extent to which the ruling groups have accepted the view that full political and economic membership of the state should be extended to all its citizens and is moving in that direction; (b) the existence of marked inequalities in the society without religious or other ideological support for such inequalities.

States which are moving towards granting more citizenship rights to all its members are likely to depend increasingly on education and less on the use of physical control mechanisms in their efforts to ensure social stability. This will be true even of societies which are characterized by gross inequalities in the distribution of wealth and status without any accepted religious or other ideological support for such inequalities. In such situations, the problem of social stability will loom larger and the state will likely depend at first on the use of physical control mechanisms to prevent social disruption. But as it moves to widen areas of citizenship to all sections of the population, its use of education in political socialization, i.e., developing acceptance of the social order, is likely to increase. Schools will attempt to instill in their students the ideals which are supposed to be embodied in the political system and to prepare them to take an active role in it or develop confidence in those who manage it. This contribution of education to such efforts is likely to be more effective if interlaced with some religious or ideological belief. Religion, which Marx described as “the opium of the masses” is often seen as an effective instrument in domesticating the demands of the “lower classes” — a condition which is usually considered a necessary prerequisite for granting them more citizenship rights.



Further, since education is often a means of upward social and economic mobility — even though its role in this direction in well established, hierarchically structured societies is often more apparent than real — it helps to nurture the general belief among the population that their society is open and its rewards are mainly distributed on the basis of one's own efforts. This gives education an important role in the containment of potential conflict and in the "cooling off" of aspirations which are unsupported by educational achievement. This point was recognized by the National Commission on Civil Disorders in the U.S.A. In its Report (1968) the Commission observed that

*Level of schooling is strongly related to participation (in riots). Those with some high school education were more likely to riot than those who had finished grade school. High levels of education and income not only prevent rioting but are more likely to lead to active responsible opposition to rioting [emphasis added]. (p. 425)*

Increasingly, as schooling spreads and as citizenship rights widen, academic and other forms of achievement become more acceptable bases for social and economic differentiation. As this happens, the dependence on religion to give "moral" support to existing societal inequalities decreases. Also, as schools succeed in inculcating those values which are supportive of the existing social order and developing a wider popular acceptance of inequalities based on achievement, the need for the use of physical mechanisms in ensuring social stability diminishes except where there are gross and obvious discrepancies between the values inculcated by the school and the realities of life in the society. Incidentally, it should be noted that the decision of a state to expand education might not reflect any desire of the rulers to extend citizenship rights to the population, but that education might be seen simply as a means of technical manpower formation.

In review one sees that, in the West Indies, as it became necessary to extend certain citizenship rights to the ex-slaves, the planters changed their attitudes to the provision of education for this group. Schools were rapidly established under the auspices of various religious groups in the immediate post-emancipation period and one of the major aims of the education offered was to get the new generation to accept some of the basic values which provided the "moral" underpinning for the "new" plantation society. These citizens of the future were taught in schools to accept the superiority of the European and his culture, to be content with their positions in life, and to be obedient citizens. As Professor Braithwaite (1967) argued, the educational goals in the West Indies during this period were based

on the incorporation of slaves as citizens, with the church being used as an agency for this purpose, and its teaching stressing one's duty to one's neighbour, the virtue of humility, the need to carry oneself lowly and reverently before one's teachers, masters and betters, to be content with one's lot and to be an obedient citizen. (p. 18)

Education gradually became very effective in helping to develop in the society a strong support for inequality which was based on achievement. In fact, the gradual decline in the importance of church control of education in the West Indies can, to some extent, be seen as related to the success which education had in developing public acceptance in the society of marked inequalities which seem to be based on the achievement of individuals or their families.

My thanks are due to my colleagues Dr. R. S. Pannu for his comments on an earlier draft of this paper and to Dr. P. Rooke from whose doctoral dissertation I have drawn heavily for some of the points raised and references used in this paper. Mrs. S. El Nahas's help in editing this paper and Anne Cooper's assistance with the typing are also gratefully acknowledged.



*References*

- Agricultural Reporter*, January, 1892, reported in Gordon.
- Augier, F. R., & Gordon, S. *Sources of West Indian history*. London: Longmans, 1962.
- Braithwaite, L. E. The school curriculum and the emergent needs of society. In *Report on Conference on Teacher Education in the Eastern Caribbean*, V. WI. Institute of Education, 1967.
- British Government. *Circular Dispatch*, October 1, 1838, reported in Gordon.
- Coleridge, H. N. *Six months in the West Indies in 1825*. London: John Murray, 1832.
- The Colonist*, February 18, 1824.
- Daly, V. T. *A short history of the Guyanese people*. Georgetown, Guyana: Daily Chronicle, 1966.
- Fanon, F. *Towards the African revolution*. New York: Grove Press Inc., 1967.
- Gordon, S. *A century of West Indian education: A source book*. London: Longmans, 1963.
- Government of Great Britain, Board of Education. *Special report on education subjects, Vol. 4: Educational systems of the chief colonies in the British Empire*, H.M.S.O., 1901.
- James, C. L. R. The slaves. In L. Comitas & D. Lowenthal (Eds.), *Slaves, freemen, citizens: West Indian perspectives*. New York: Anchor Books, 1973.
- Latrobe, C. *Report on the Windwards and the Leewards*, 1838, reported in Gordon.
- Martin, R. M. *The British colonies, their history, extent, conditions and resources*. London: J & F Tallis, Vol. 8, undated.
- Massialas, B. *Education and the political system*. Reading, Mass.: Addison-Wesley Publishing Co., 1969.
- Moore, W. E. *Social change*. New York: Prentice Hall Inc., 1963.
- Myrdal, A. The principle of cumulation. In A. Etzioni & E. Etzioni, (Eds.), *Social change*. New York: Basic Books Inc., 1964.
- Patterson, O. *The sociology of slavery: An analysis of the origins, development and structure of Negro slave society in Jamaica*. London: McGibbon & Kee, 1967.
- Report of the National Advisory Committee on Civil Disorders. New York: Bantam Books Inc., 1968.
- Rooke, P. T. *The Christianization and education of slaves and apprentices in the British West Indies*. Ph.D. thesis, The University of Alberta, 1977, (unpublished).
- Secretary of State for the Colonies. *Circular dispatch to governors*, 15th October, 1835.
- Smith, Rev. J. *Communication: Smith to Burder*, March 18th, 1818 (Box 2, Br. G/D 1815-27), London Missionary Society Archives.
- Smith, R. T. *British Guiana*. London: Oxford University Press, 1962.
- Smith, R. T. Social stratification, cultural pluralism and integration in West Indian societies. In S. Lewis & T. G. Matthews, (Eds.), *Caribbean integration*, Rio Pedras: University of Puerto Rico, 1967, p. 234.
- Sterling, Rev. J. Report to the British Government, 11th May 1935. See Gordon.
- Waddell, H. M. *Twenty-nine years in the West Indies and Central Africa. A review of missionary work and adventure*. London, 1863.

RUTH BRAGMAN

and

ROBERT HARDY

*University of Maryland*

## Identical and Reverse Visual Pattern Recognition in Deaf Children

*Twenty deaf children were presented with a pattern recognition task. The purpose was to investigate the development of pattern recognition and pattern reversal in the deaf child aged six through eight, and its relation to age of exposure to a gestural symbol system. Results indicated that (1) same pattern recognition was significantly higher than reverse pattern recognition, (2) age of exposure to a gestural symbol system did not affect performance on same pattern or reverse pattern recognition, and (3) age affected performance on same pattern recognition but not on reverse pattern recognition. Results are discussed in terms of a developmental hierarchy in pattern recognition tasks. (Ms. Bragman has worked with children with a variety of handicapping conditions and is presently supervising practicum students in Special Education; Dr. Hardy is a Professor in Human Development at the University of Maryland, College Park, Md.)*

The ability to recognize patterns, according to numerous studies, is seen as a necessary component for processing information and the development of cognitive processes (Anderson, 1975; Bower, 1975; Flavell, 1963, Piaget & Inhelder, 1956). Pattern recognition has been related to the development of mathematics (Davis, 1972; Lovell, 1971; Piaget, 1972; Sternberg, 1975); language development (Sternberg, 1975); problem solving (Gagné, 1970); knowledge component in learning (Bloom, 1956); and the elements of intelligence (Guilford, 1967). The ability to recognize patterns has also been used as a component of many intelligence tests as a means of assessing a child's cognitive functioning (Leiter, 1959; Raven, 1956; Thurston, 1938; Weschler, 1967). Research in pattern recognition tends to show that there is a hierarchy in cognitive ability beginning with the ability to match patterns by the age of four and five, and the ability to display reversals by the age of six and seven (Flavell, 1963; Piaget & Inhelder, 1956; Sternberg & Larson, 1976).

Sternberg and Larson (1976), developed a hierarchy in pattern recognition for

normal children considering the aspects of pattern sequence, stimulus property, and pattern task. In a study of 96 subjects from pre-kindergarten to first grade, they noted there was a significant difference between subjects due to grade, ability and task. More significantly, an interaction was found between grade level, ability level, and pattern task, indicating support for Piaget's contention that cognitive development is not simply a grade-level or ability-level phenomenon.

The literature on the cognitive development of hearing impaired children addresses the issue of the role of language in cognitive development. Furth (1973) extends Piaget's theory of development and emphasizes that development of cognitive processes of the deaf is the same as in hearing children. He feels that the reason most deaf children perform poorly on the tasks is due to the linguistic quality of the tasks asked of them. Furth (1973), in a study of a group of deaf children, six to fourteen years of age, and hearing children of the same ages, found that there was no consistent differences between deaf and hearing children on symmetry or sameness. It was noted, however, that there were fairly large differences in favour of hearing children on opposites. Furth suggested this was due to some extent to the limitations presented by sign language in expressing opposites, and felt that there may be a slight lag in deaf subjects of one to two years. This differed from previously reported lags of five to ten years. Furth also maintained that deaf children exhibited the same capacity as hearing children in inference, if the linguistic component was eliminated. However, Suppes (1974) questions this assumption regarding the inference test, since even the simple inference tasks given young normal children (aged six and seven) are too difficult to test outside of a verbal context.

Bert and Roberts (1976) have studied the early cognitive development of hearing impaired children. In their study involving 16 deaf children, it was shown that subjects progressed normally through the period of sensorimotor development, except in the area of vocal limitation. Tomlinson-Keasey and Kelly (1974) also felt that the early cognitive development of deaf children was normal; however, they felt that from 18 months to four to five years of age, the child becomes frustrated if he lacks models for a well-developed gestural symbol system which would enable him to communicate.

The need for an early gestural symbol system is supported by much of the research. Vernon and Koh (1970) matched 32 pairs of genetically deaf children for sex, age, and intelligence; one group, exposed to manual communication from infancy, was higher by 1.44 years. Meadow (1976) compared 59 deaf children of deaf parents with deaf children of hearing parents and also reported that children with deaf parents ranked higher in academic achievement.

The purpose of this study was to investigate the development of pattern recognition and pattern reversal in the deaf child aged six through eight and its relation to age of exposure to a gestural symbol system.

The findings of Sternberg and Larson (1976) indicate that there is a hierarchy in pattern recognition skills and provide support for the developmental theory of Piaget. They present evidence that the hierarchy in pattern recognition skills begins with same pattern recognition, then develops to reversal pattern recognition, followed by intradimensional shift and finally extradimensional shift. This study attempted to support the assumption that since same pattern recognition develops before reversal pattern recognition, all deaf children, aged six



through eight, would perform better on same pattern recognition than on reversal pattern recognition.

The research of Vernon and Koh (1970) and Meadow (1956) suggested that children who have been exposed to a manual communication system from infancy showed higher academic achievement, of 1.44 years, than those children not exposed to a manual communication system in infancy. Sternberg and Larson (1976) findings indicated that there is a significant performance increase on same pattern recognition at five years of age, though improvement on reverse pattern recognition does not occur until seven years of age. If children not exposed to manual communication at infancy are at least a year behind children exposed to manual communication at infancy, one would expect that the ability to perform reverse pattern recognition would not be as well developed in children not exposed to a gestural symbol system at an early age. This study predicted that the earlier the age of exposure to symbol system the better the performance.

The studies of Sternberg and Larson (1976), Vernon and Koh (1970), and Meadow (1956) reported an interaction between grade level, ability, and pattern task. These findings showed that pattern recognition performance correlates with grade level and ability in combination, rather than with these two factors in isolation. Based on these findings, it was predicted that older children would perform better than younger children on same pattern and reverse pattern recognition.

From the above discussion, three null hypotheses were postulated:

1. For all deaf children, there is no difference in performance in pattern recognition between same pattern and reverse pattern.
2. There is no difference in the performance of the deaf child due to age of exposure to a gestural symbol system on total pattern recognition, same pattern recognition, and reverse pattern recognition.
3. There is no difference in the performance of the deaf child due to age on the total pattern recognition, same pattern recognition, and reverse pattern recognition.

### *Method*

#### *Subjects*

Twenty subjects were selected from a total communication deaf school in suburban Maryland; they ranged in age from six years, four months to eight years, nine months. All subjects had been deaf since before the age of three; age of exposure to a gestural symbol system ranged from infancy through four years.

#### *The instrument*

The instrument consisted of 20 cards, 14 by 22 inches (35.56 cm by 55.88 cm), 4-ply, white poster board, bearing one model sequence pattern and four choice sequence patterns. Each pattern was made up of 1-inch (2.54 cm) squares of coloured paper as the dimension. On the left side of each card the model pattern sequence was given, represented by the coloured squares. Four patterns were used (ABAB, ABCABC, ABBABB, AABAAB). On the right side of the card four choice pattern sequences were presented. On each card one of the four choice pattern sequences was the same as the model pattern sequence, one was the reverse of the model pattern sequence, and the other two were randomly selected patterns that

included only those colours given in the model sequence. All choice sequences were randomly placed on the cards. Cards were randomly ordered.

### *Treatment*

Each subject was first given the Ishihara Color Blind Test (Ishihara, 1970). No subjects were colour blind. Each subject was then given three practice trials on the pattern recognition tasks. An interpreter was present at all times and gave the following directions:

Look at these colours (pointing from left to right at the model sequence pattern), now look at these colours (pointing, one at a time, from left to right at choice sequence patterns). Point to the one that is the same (or reverse of, depending on the treatment group) as this one (pointing to the model sequence pattern).

If a subject failed to point to a pattern in the choice sequence patterns, the directions were repeated and the correct answer was shown. If after three repetitions of directions the subject still failed to point, the subject was eliminated from the study and another subject was selected from the subject pool. One subject had to be replaced.

Between each treatment there was a five-minute break, and three new practice trials were given before the second treatment was administered. Testing took 30 minutes per subject.

### *Treatment Groups*

All subjects received two treatments. Treatment one was the pattern task of finding the identical pattern match. Treatment two was the pattern task of finding the reverse pattern match. Subjects were randomly assigned to a balanced order of treatment.

## *Results*

### *Hypothesis 1*

The means and standard deviations for the same pattern recognition task were 18.85 and 1.76 respectively, and for the reverse pattern recognition task, 15.15 and 6.94 respectively. The higher the score the better the performance, with the highest possible score being 20. To test the difference between these means, a repeated measure design was run. Table 1 presents a summary of this analysis.

The results indicate that performance on the same pattern recognition task was significantly higher than performance on the reverse pattern recognition tasks [ $F(1, 19) = 4.93, p < .05$ ]. Thus the first null hypothesis that there is no difference for all children in performance between same and reverse pattern recognition was rejected.

### *Hypothesis 2*

Hypothesis 2 was tested by performing a multiple regression analysis to eliminate the effect of order of tasks on performance first, and then to analyze the effect of age of exposure of a gestural symbol system. The multiple  $R$ s for effects of age of exposure to a gestural symbol system on total pattern recognition, same pattern recognition, and reverse pattern recognition were .3899, .1178, and .3603 respectively.

TABLE 1  
REPEATED MEASURES ANALYSES OF SAME AND  
REVERSE PATTERN RECOGNITION

Source	SS	df	MS	<i>F</i> *
Task	136.90	1	136.00	4.93**
Subjects	446.00	19	23.47	
Interaction	427.10	19		
Total	1010.00			

\* *F* of 4.35 needed for significance beyond the .05 level.  
\*\* significant, *p* < .05

TABLE 2  
ANALYSIS OF EFFECTS OF AGE OF EXPOSURE TO GESTURAL SYMBOL SYSTEM  
ON PERFORMANCE ON PATTERN RECOGNITION TASKS

Source	SS	df	MS	<i>F</i> *
<u>Total pattern recognition</u>				
Order	72.20	1	72.20	1.62
Age of exposure	63.41	1	63.41	1.43
Error	756.40	17	44.50	
Total	892.00	19		
<u>Same pattern recognition</u>				
Order	.05	1	.05	.01
Age of exposure	.76	1	.76	.22
Error	57.74	17		
Total	58.55	19		
<u>Reverse pattern recognition</u>				
Order	68.45	1	68.45	1.36
Age of exposure	50.27	1	50.27	1.07
Error	795.83	17	46.81	
Total	914.55	19		

\* *F* of 4.45 required for significance beyond the .05 level.

An examination of these results indicates that age of exposure to a gestural system does not affect performance on any of the tasks. Thus, the second hypothesis that there is no difference in the performance of deaf children due to age of exposure of a gestural symbol system on total pattern recognition, same pattern recognition, and reverse pattern recognition was not rejected.

*Hypothesis 3*

Hypothesis 3 was tested by performing a multiple regression analysis to eliminate the effect or order of tasks on performance first, and then to analyze the



TABLE 3  
ANALYSIS OF EFFECTS OF AGE ON PERFORMANCE  
ON PATTERN RECOGNITION TASKS

Source	SS	df	MS	F*
<u>Total pattern recognition</u>				
Order	72.20	1	72.20	1.58
Age	40.74	1	40.74	.89
Error	779.06	17	45.83	
Total	892.00	19		
<u>Same pattern recognition</u>				
Order	.05	1	.05	.02
Age	13.19	1	13.19	4.91**
Error	45.31	17	2.67	
Total	58.55	19		
<u>Reverse pattern recognition</u>				
Order	68.45	1	68.45	1.56
Age	100.31	1	100.31	2.29
Error	745.80	17	43.87	
Total	914.55	19		

\* *F* of 4.45 required for significance beyond the .05 level.

\*\* significant,  $p < .05$

effects of age. The multiple *R*s for effect of age on total pattern recognition, same pattern recognition, and reverse pattern recognition were .3558, .4756, and .4296 respectively. Table 3 presents a summary of these results.

An examination of these results indicates that age affects the performance on the same pattern recognition, but does not affect the performance on either total pattern recognition or reverse pattern recognition. Thus, the third hypothesis that there is no difference in the performance of deaf children due to age on total pattern recognition, same pattern recognition, and reverse pattern recognition cannot be completely rejected.

### Discussion

Since all children performed better on same pattern recognition than reverse pattern recognition, the first null hypothesis was rejected. This supports the previously mentioned position of Sternberg and Larson (1976) that there is a hierarchy in pattern recognition skills and provides support for the developmental theory of Piaget.

Since there was no effect in pattern recognition performance on same pattern recognition, reverse pattern recognition, and total pattern recognition due to age of exposure to a gestural symbol system, the second null hypothesis could not be rejected. These results may be explained in either of two ways. Since mean age of exposure to a symbol system was 1 year 9.5 months, this early exposure to a symbol system may have affected the findings. This is supported by the findings of Tomlinson-Keasey and Kelly (1974) and Bert and Roberts (1976) that early cognitive development of hearing impaired children is normal but there is a need

for an early gestural symbol system between the ages of 18 months to four or five years. None of the children in this study were exposed to a symbol system later than four years of age. This lack of relationship between the age of exposure to a gestural symbol system and pattern recognition tasks can also be further support to Furth's (1973) contention that cognitive development is independent of language.

The age of the subjects affected only the performance on the same pattern recognition but not on the total pattern or reverse patterns. Thus, the third null hypothesis was not completely rejected. This result may also be explained in either of two ways. Since the mean age of the group was seven years, 8.3 months, which is above the age at which Sternberg and Larson (1976) found the skill of reversals to appear, there is the possibility that ability may have confounded the findings. These results, however, can also be interpreted in respect to Furth's (1973) suggestion that the reason for the deaf child's lower performance on opposites is due to the linguistic component of the task. While the task of identical may not have this linguistic component the reverse task may have; therefore, the children's performance was lower on this task due to this component.

This difference in performance becomes clearer when observations were made of the individual performances of the children. Eighty percent of the children were able to perform the same pattern recognition task to 90% competency while only 60% of the children were able to perform the reverse pattern recognition task to 90% competency. Ninety-six percent of the wrong answers given on the reverse pattern recognition task were that of an identical pattern match. By observation of the children during the reverse pattern recognition task, it was noted that all the children could do the specified task on the practice trial; however, they failed to do the task as specified on the actual task. The children would look (or point as in two cases) at the model pattern from right to left and then look at each of the choice patterns from right to left, thereby choosing the identical pattern instead of the reverse pattern. This leads one to question if this is a step in the development of the concept of reverse or the effect of the linguistic component of the concept of reverse, since the children did indeed reverse the patterns.

Further research is indicated in the area of the development of reverse pattern recognition across various populations. This would include children from various language backgrounds that may have different linguistic components involved in the concept of reverse to determine if this is a developmental stage or a linguistic problem.

#### References

- Anderson, B. F. *Cognitive psychology*. New York: Academic Press, 1975.
- Bert, B., & Roberts, G. Early cognitive development in hearing impaired children. *American Annals of the Deaf*, 1976, 121, 560-564.
- Bloom, B. *Taxonomy of educational objectives. Handbook I: Cognitive domain*. New York: David McKay Co., 1956.
- Bower, G. Cognitive psychology: An introduction. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. New York: John Wiley and Son, 1975.
- Davis, R. Naive foundations for a theory of mathematics learning. In W. E. Lamon (Ed.), *Learning and the nature of mathematics*. Palo Alto: Science Research Associates, Inc., 1972.
- Flavell, J. *The developmental psychology of Jean Piaget*. New York: Van Nostrand Reinhold, 1963.

- Furth, H. G. *Deafness and learning*. Belmont, California: Wadsworth Publishing Company, Inc., 1973.
- Gagné, R. M. *The conditions of learning*. New York: Holt, Rinehart and Winston, 1970.
- Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.
- Ishihara, *Ishihara Color Blind Test Book*. Tokyo: Hanehara Shuppan, 1970.
- Leiter, R. G. *Leiter International Performance Scale*. Chicago: Stoetling Co., 1959.
- Lovell, K. *The growth of understanding in mathematics*. New York: Holt, Rinehart and Winston, Inc., 1971.
- Meadow, K. The effect of early manual communication and family climate on the deaf child's development. In L. Mann & D. A. Sabatino (Eds.), *The third review of special education*. New York: Grune and Stratton, 1976.
- Piaget, J. Mathematical structures and the operational structures of intellects. In W. E. Lamon (Ed.), *Learning and the nature of mathematics*. Palo Alto: Science Research Associates, Inc., 1972.
- Piaget, J., & Inhelder, B. *The child's conception of space*. London: Routledge and Kegan Paul, 1956.
- Raven, J. *Coloured Progressive Matrices*. London: L. K. Lewis, 1956.
- Sternberg, L. Pattern recognition training: A key to mathematics and language skills development. *Teaching Exceptional Children*, 1975, 7, 61-63.
- Sternberg, L., Epstein, M. H., & Adams, D. Performance characteristics of retarded and normal students on pattern recognition tasks. *Contemporary Educational Psychology*, 1977, 2, 209-218.
- Sternberg, L., & Larson, P. The development of pattern recognition ability in children. *Contemporary Educational Psychology*, 1976, 1, 146-156.
- Suppes, P. A survey of cognition in handicapped children. *Review of Educational Research*, 1974, 74, 145-176.
- Thurston, L. L. *Primary mental abilities*. Chicago: University of Chicago Press, 1938.
- Tomlinson-Keasey, C., & Kelly, R. R. The development of thought processes in deaf children. *American Annals of the Deaf*, 1974, 119, 693-700.
- Vernon, M., & Koh, S. Early manual communication and deaf children's achievement. *American Annals of the Deaf*, 1970, 115, 527-536.
- Weschler, D. *Weschler Intelligence Scale for Children*. New York: The Psychological Corp., 1967.



WERNER LIEDTKE

and

GWEN STOTT

University of Victoria

## Games and Game Settings for the Preschool Child

*The ideas for game settings and the general comments in the article were selected from the summaries of two observational projects. For one setting, games were designed. These games had one common element — there existed the possibility for the use of some sort of strategy. The following sample questions illustrate the main focus of attention as game playing behaviour was observed: Will young children be able to give reasons for their own or their opponents' moves? Will they modify their "strategy"? Will they copy an opponent? Can they give reasons for a move made by an opponent? For another setting, unfamiliar and familiar game-boards were presented to the children who were asked to invent their own games and rules, or to describe how they thought a particular game was to be played. (Mr. Liedtke is an Associate Professor in the Faculty of Education, and Ms. Stott is an elementary teacher in Leduc, Alberta.)*

The statement, "Let's play a game" is more than likely to arouse a favourable reaction from most children. Even preschool children will show enthusiasm and excitement when the word game is mentioned. One five-year-old summed up how he felt about playing games by stating emphatically, "Yeah, that was neat. I never get tired of games — no way." Her comment illustrates how most of the four-, five-, and six-year-olds felt who took part in two different projects<sup>1</sup> that involved a variety of games. As part of these projects, the young children were observed in different game settings. The children were given an opportunity to react to the rules they were using. They also had a chance to modify existing rules and to create rules and games of their own. Of all the different observations that were made during the projects, one result stands out above all others. Given an opportunity, young children are all too eager to participate. They enjoy the involvement, ideas are exchanged freely among peers and with adults in charge, and a lot of satisfaction seems to be derived by the children from being able to present their own ideas or suggestions.

In the paragraphs that follow, some of the games and game settings that were used during the aforementioned projects are described and illustrated. It is hoped that the reader can make use of some of these games and ideas either as a

supplement to an existing collection of games, as a beginning to a collection suitable for young children, or for settings where young children's behaviour and/or thinking is observed.

Several of the game settings involved advancing a marker on a gameboard by matching steps to be taken by this marker with the number of dots shown on a homemade number cube or die (sponge cube with felt markings). Teachers of young children know that this matching procedure can be a difficult task for some children. The most common mistakes include counting the space occupied by the marker as the first step, assigning more than one number to a step, and skipping a space while counting steps. One activity and one game setting that can be used to assist young children who make these mistakes consist of the following. The children take turns rolling a number cube (i.e., 

·	·	·	·	·	·
---	---	---	---	---	---

). After each roll, the children are asked to hold up as many fingers as there are dots and to take as many steps as there are fingers or dots. For the game, two or more children line up against a wall and take turns rolling or dropping a number cube. The child who has rolled the number cube holds up as many fingers and takes as many steps away from the wall as there are dots on the cube. The child who is the first to get across the room to the opposite wall is declared the winner. It is fascinating to observe the development of "step-taking strategy" as some four-year-old children attempt to play this game of "Race Across the Room." It may take as many as three or four games and a few appropriate comments from the opponent, such as "I won again," or "I think I won because I take big steps," before these young children realize that it is advantageous to take "larger" steps. However, more than one four-year-old child was encountered who was not moved by either comments or by losing a game. The playing style was not changed; small steps continued to be taken.

During the projects an attempt was made to accommodate different ability and age levels for various game settings. To accomplish this, slight variations were introduced. For example, the changes that were made for the game "Race Across the Room" included: showing greater numbers on the cubes, showing different arrangements for numbers, using number names or numerals, including zero as one of the possibilities on the cube, and using numbers and number names.



Figure 1

Any setting that involves rolling a number cube or die and advancing a marker along a track of some sort can be a learning experience for a young child, since the strategies of matching and counting are involved. Some five-year-old children are ready for the variation that involves rolling two number cubes and then advancing a marker according to the number of dots shown on both of these cubes. In this setting, the children should be allowed to use and/or develop their own strategies for finding the total after each roll. If that is the case, different methods of solving the problem will surface and the children who are observing each other will become aware of the fact that there exist more than one way of finding the answer.

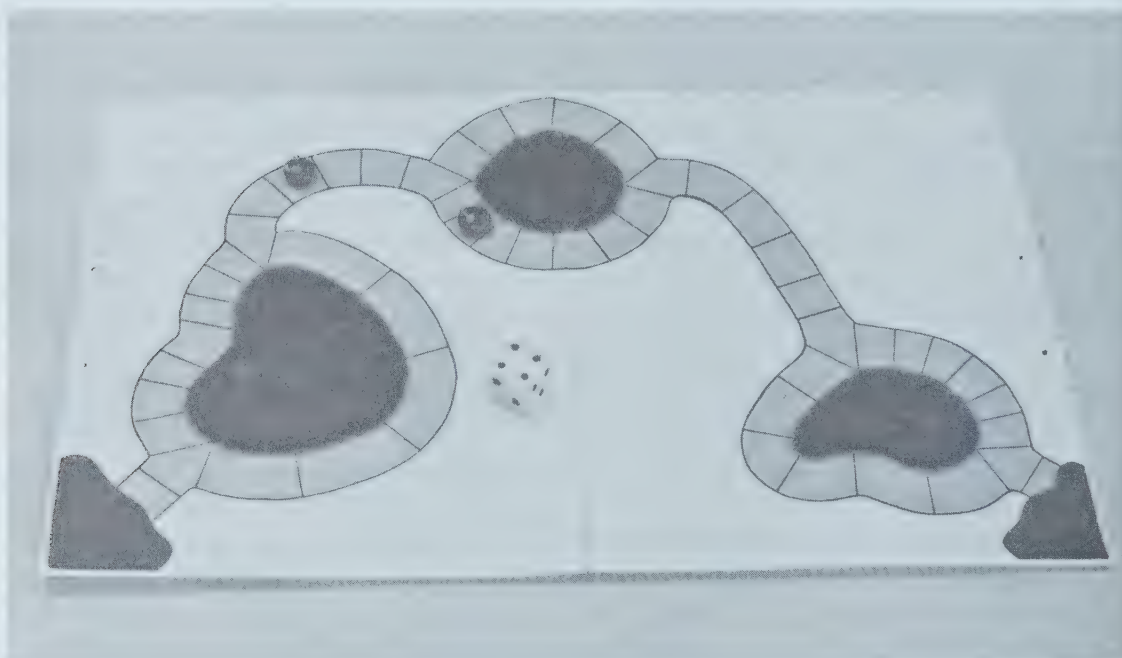


Figure 2

The strategies of matching and counting are involved in the "Frog Race" setting illustrated in Figure 2. For this game, a number cube is rolled, a marker (frog) is advanced, and the winner is the child whose frog is the first to get to the other side of the gameboard. However, as part of the game setting, young children are given an opportunity to make decisions. As they advance a marker across the board, they may choose between routes. The setting is such that the longer looking path is actually shorter, or vice versa, or two paths that look to be the same are actually different. As part of one project, the children were asked to verbalize their reasons for taking a certain route. The majority of responses indicated, as one would expect, that young children rely heavily on perception and make decisions accordingly. Two simple variations for this game setting consist of reducing the number of choices or "grassy spots" or substituting number cubes with different markings (Figure 1).

A geoboard was used to create the game setting shown in Figure 3. This game was presented to the older children who participated. The eight possible directions ( $\uparrow \downarrow \rightarrow \leftarrow \nearrow \nwarrow \swarrow \searrow$ ) that can be taken on a geoboard are drawn on small pieces of cardboard (i.e.,  $\uparrow$ ). Six cards for each direction are prepared. Each of two players is assigned a different coloured bead. For the game, these beads are placed in the centre of the board. The pieces of cardboard showing the different directions are turned face down. The players take turns lifting one card. After a player turns a



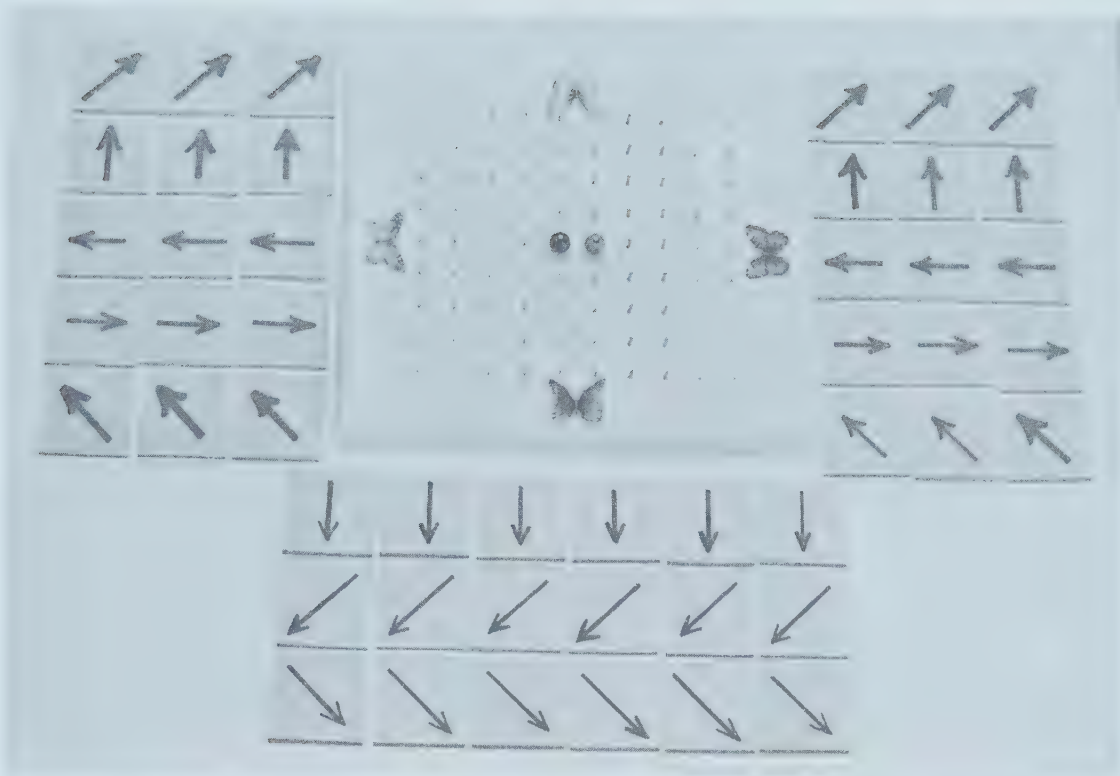


Figure 3

card, his bead is moved onto the next nail, as indicated by the directional arrow on the card. The first player to get to a nail near a butterfly is the winner.

To give young children an opportunity to make a decision or to choose between alternatives, the following variation was introduced. For each turn, the players are permitted to turn over two (or three) cards. A choice is then made in terms of the direction a player thinks is most advantageous for racing to a designated butterfly. Another simple variation consists of reducing the possible directions to  $\uparrow$   $\downarrow$   $\rightarrow$  and

A simple tic-tac-toe setting is illustrated in Figure 4. A colour is chosen and then turns are taken in placing the coloured corks into the holes on the board. The first player to get four corks in a row, vertically, horizontally, or diagonally, is the winner. The children seemed to enjoy this game. The majority were able to think only in terms of offensive moves, though several of the older children showed awareness of both offensive and defensive strategies. One six-year-old offered the following "valuable" advice to her opponent who had just lost the second game in a row, "You got to be really smart. You got to think that I might win if I'm going this way — this way — then there — and if I only got one more to go, you got to block me." (This game setting can easily be duplicated with coloured beads and a geoboard.)

One game which is interesting in terms of observing a variety of reactions, comments or facial expressions is "Tic-tac-toe — Three in a row!" For this game, the adult takes the first turn in placing a cork somewhere near the middle of the gameboard. Then turns are taken and the winner is the player who is first to get three corks in a row, vertically, horizontally, or diagonally. Some young children will play the game twice and realize right away what is going on. Others play it four or more times before the expression on their faces shows that they have made a

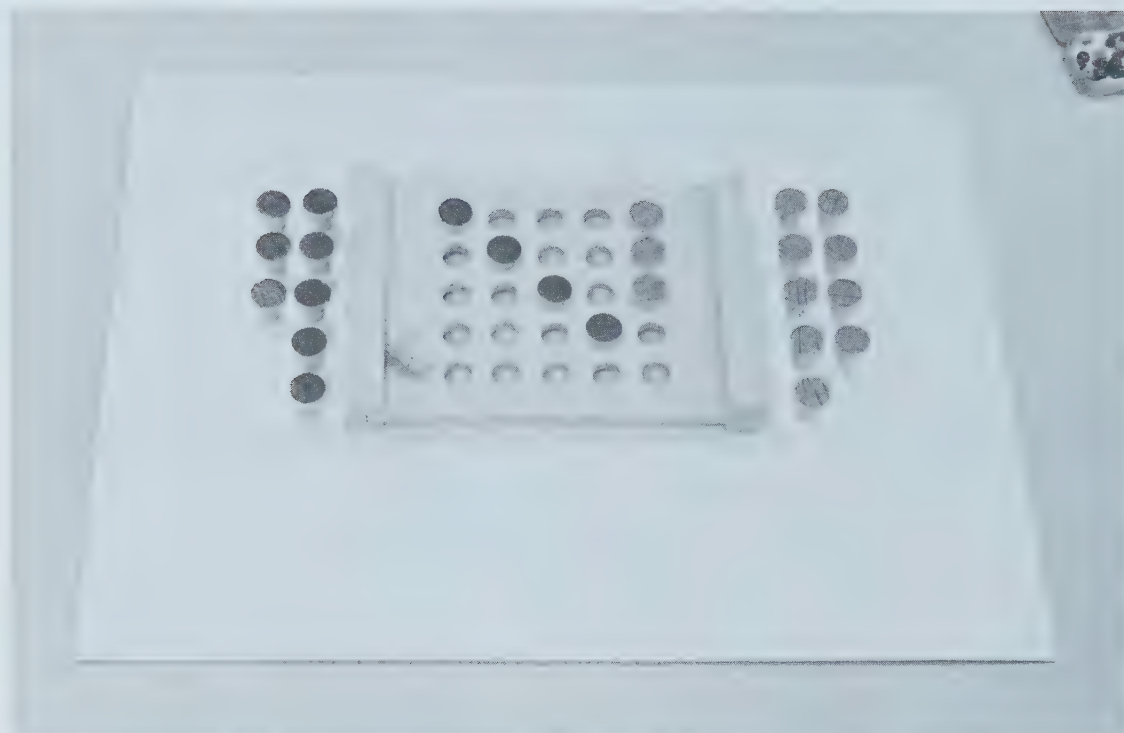


Figure 4

deduction about the game. The comment, "I don't like this game," is common and the majority of children realize that "I can win if you let me start."

For the "Farm-Animal-Race" (Figure 5), players take turns throwing a number cube. Animals are advanced along a path in an attempt to be the first across the board. Along the way, the animals encounter red and/or yellow spots on the gameboard which require moving ahead or back depending on the direction indicated by an arrow. After the children are familiar with the setting, the

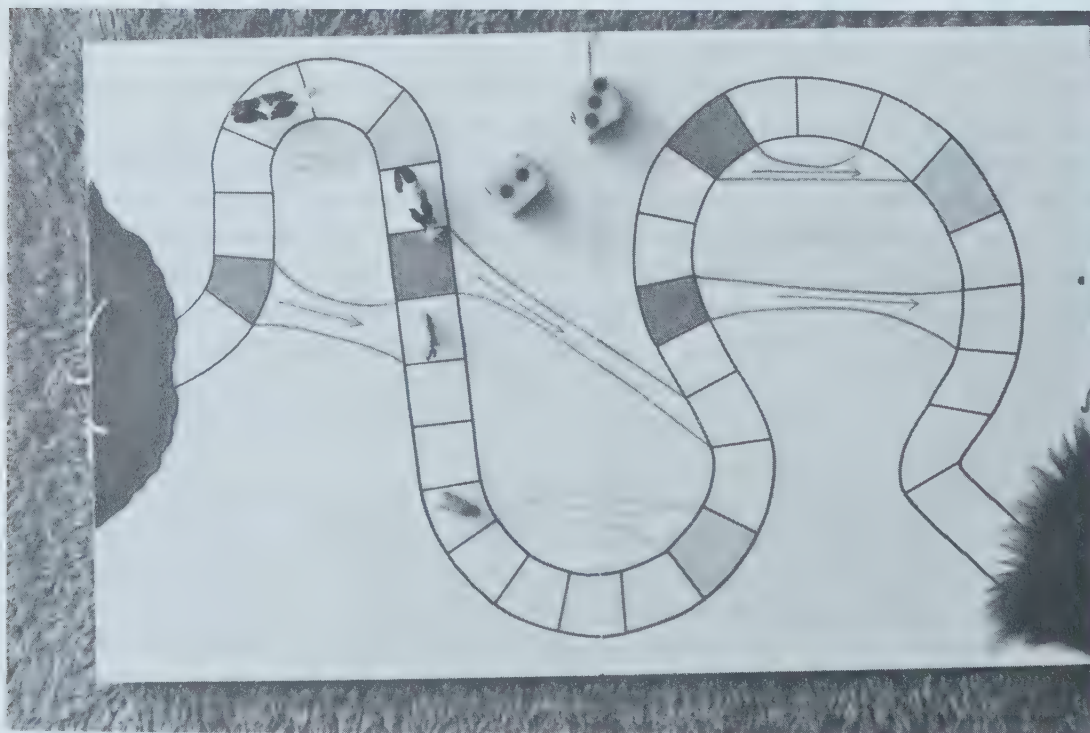


Figure 5



following variation is introduced. Each player is assigned two animals and both are to be advanced across the board, simultaneously or separately. The reason for this variation was an attempt to find out whether or not young children would or could develop the strategy of dividing up the steps to be advanced between the two animals in order to avoid sliding back and/or in order to take greater advantage of the spots that enabled them to move ahead. For the younger children, there were too many variables in this setting and they advanced one marker at a time. One child explained, "I can't manage two at a time." A few of the six-year-old children soon discovered the intended strategy. One child evaluated the game by stating his reason for enjoying the game, "Because you get to slide — slide — slide — all over the place!"

At times the gameboards and materials described above were presented without explanation of rules and different children were asked to think up possible uses for these. Inventing games and creating rules were enjoyable experiences for the children. Rules to accommodate all kinds of situations were devised and reasons for these rules willingly verbalized in what often amounted to very detailed descriptions. At times, quite a few rules were made up and it was amazing to observe that the children who had invented and described these rules remembered them as the games were played. Frequently, rules were revised or added to while a game was in progress. A revision usually occurred when children found themselves in a disadvantageous position on the gameboard. Although the reason for a revision could be observed, it was seldom verbalized by a child. In general the improvisations included the predictable setting where children attempted to transfer rules and moves from a familiar game to the gameboard on hand. However, as is the case with young children, there is always the unpredictable. The following examples illustrate different types of improvisations.

Two boys used the geoboard-butterfly setting (Figure 3) for a game that involved racing around the outside nails on the board to see who could be first to visit all four butterflies. A number cube was used to determine how many nails a bead was to be advanced.

The suggestions from the children for the tic-tac-toe materials (Figure 4) often involved "mad-scramble" games. For one such setting, two players had to put their corks into the holes as quickly as they could. Since both players were doing this at the same time, the winner was the one who got rid of the corks first.

One of the more unusual suggestions was made for the "Farm Animal Race" setting (Figure 5). One five-year-old thought the gameboard could be attached to the body with string and be worn as an advertising sign. According to the child, the sign would be an advertisement for game stores and it would "tell you what's in the shops."

Another part of the projects consisted of presenting young children with unfamiliar gameboards. (For examples, see Figures 6, 7, 8 and 9.) Number cubes, coloured chips, plastic animals and felt pens were also made available. The children were asked to make up their own games. One such gameboard (Figure 6) consists of a simple path. Happy and sad faces ("mad," according to most young children) are randomly placed on some of the subdivisions of the path. A similar but simpler setting (Figure 7) consists of a path without subdivisions. The gameboards shown in the pictures were laminated. The children were invited to use the water soluble felt pens to record any ideas for modifications on the boards.



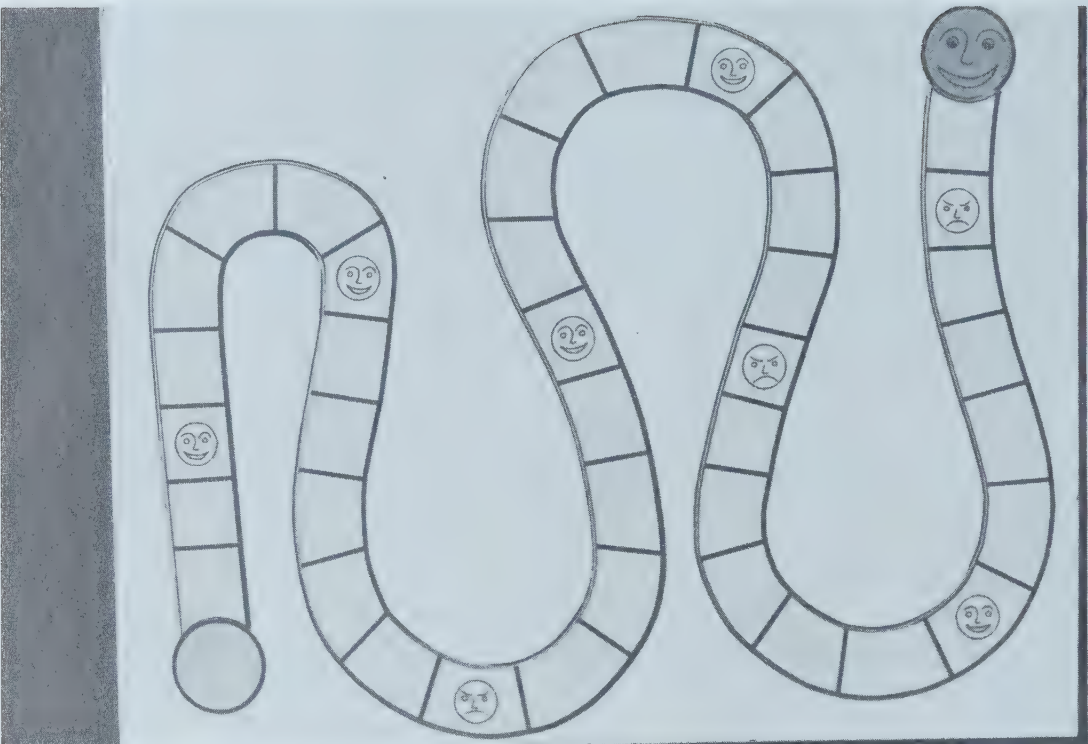


Figure 6

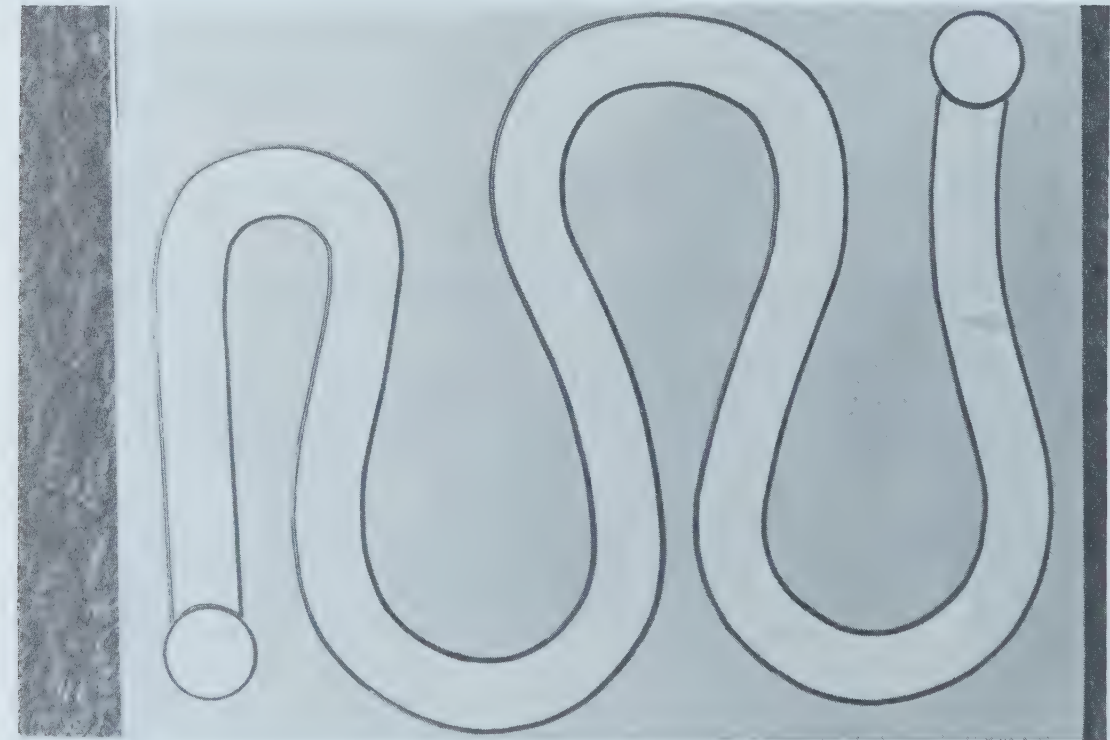


Figure 7

The children's reactions to these gameboards were diverse and very innovative, indeed. The majority of children were not disturbed by a path without subdivisions. Some of these children rolled a number cube to determine the number of steps to be taken and then took what they defined to be "middle-size" steps. A few children drew in subdivisions with the felt pens. Several children placed colored chips along the blank path to designate the steps that were to be

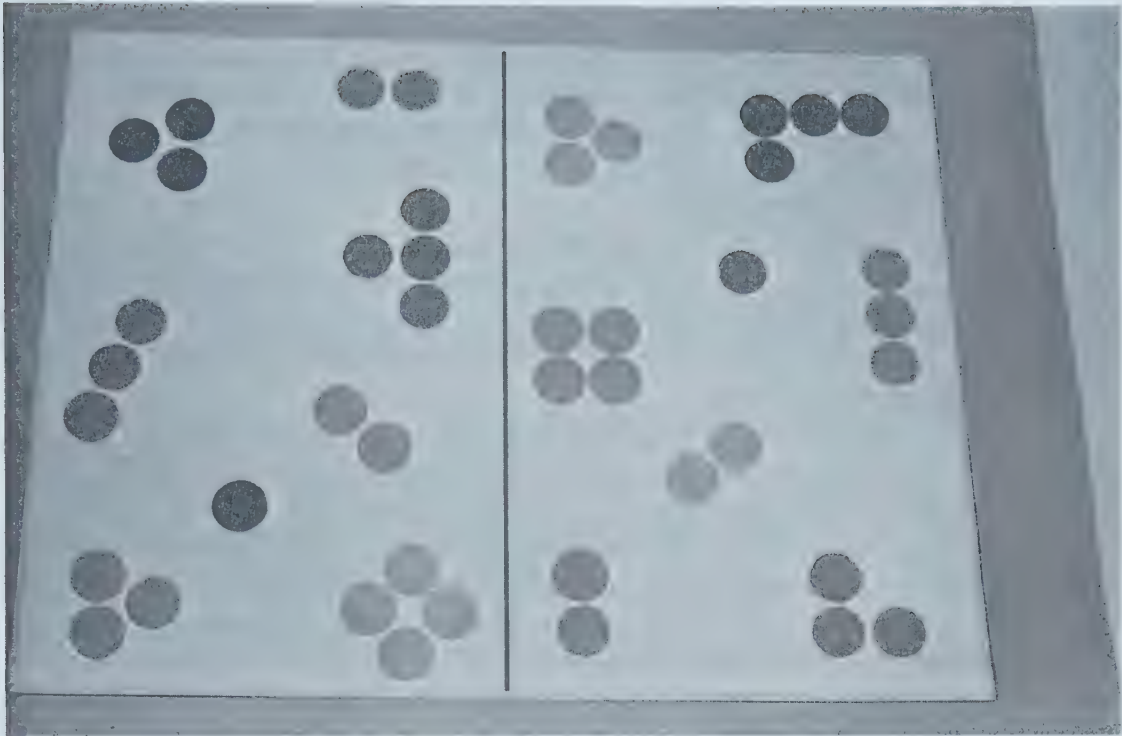


Figure 8

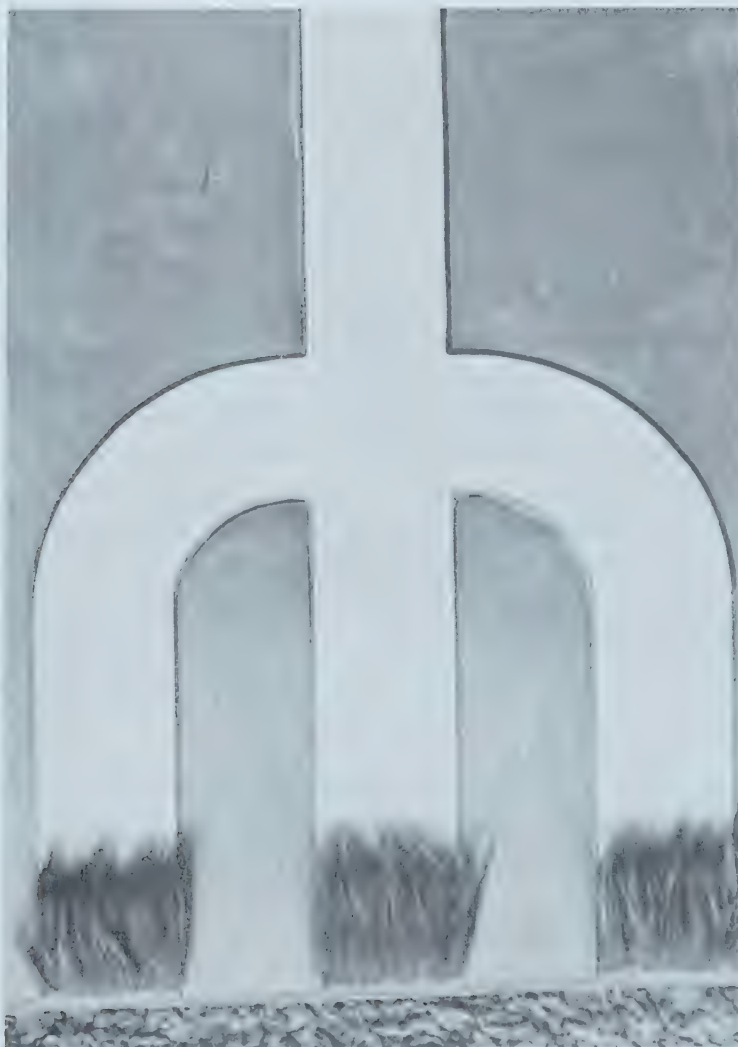


Figure 9

taken during a game. One girl devised an ingenious way to determine the number of steps to be taken without the use of a number cube. She placed six coloured plastic chips in her hand, raised her hand above the board and released some of the chips. The rule that not all of the chips are to be dropped was added to the procedure. The number of dropped chips indicated the number of steps to be taken with the marker on the gameboard.

One of the most satisfying observations included the fact that all of the children who participated in this part of the project were very eager to make up their own games. The children especially enjoyed being able to teach the adult in charge how to play their game. Even shy and quiet children talked enthusiastically about their own games and rules.

In part, the observed willingness and enthusiasm led to the following project activity. A survey of children was conducted to identify games that are frequently found in homes. The resulting list of games included Chess, Checkers, Backgammon and Master Mind. These games were obtained and presented to the young children in the following settings. If children had heard of the game and they thought they knew how to play it, they were asked to teach the game or explain the rules to the adult. If, on the other hand, the children did not know how to play the game, they were asked to suggest possible uses for the gameboards and to make up their own rules for these settings.

The observations made in this setting showed that not only did the children enjoy this setting, many of them felt proud about being able to touch and play with "adult" games. The comments the children made about these games showed that they are very observant. Many children knew quite a bit about certain game procedures and rules. However, more often than not, the rules were applied incorrectly and/or important rules were forgotten or left out when young children who thought they know how to play the game explained it to an adult. As was the case for the previous setting, the children who participated were very enthusiastic about making up their own rules for a given gameboard. More than once a child was encountered who did get carried away by his/her enthusiasm and ability to invent rules. Rules which only the creator could understand and remember were made up. However, most of these complicated games had one common element — the child who had made up the rules could not lose the game.

Any project that involves the observation of young children in action is a rewarding experience. Information can be gathered about how young children think and there are always incidents encountered which illustrate that young children differ in their thinking from adults. Learning can be observed. It happened more than once, for example, that younger children modified their game-playing responses as a result of observing older children in action. Observations can be made which reinforce what is known about young children and their behaviour. The majority of four-year-old children preferred to examine the game materials in a free or guided play setting. If possible, these children used the settings to create something from their experience. Their conversations centered around familiar objects. Five- and six-year-old children were more likely to show an interest in structured play or game settings. Most of the children in these age groups were able to justify their game moves. Many of them modified their use of strategy as a result of having observed the opponent's moves and strategy. Although their explanations for certain moves on a gameboard were often based on arbitrary terms, such as personal preference, they were beginning to understand



cause and effect relationships in game play. These older children also showed more of an interest in modifying original game rules and in creating new ones. All of the children who participated seemed to thrive on being given the opportunity to express their ideas and thoughts about games.

*Note*

1. Provincial Youth Programmes: May to August, 1977 and 1978:
    - “Young Children and Game Playing Behavior: Strategy” ( $N = 59$ )
    - “Young Children’s Thinking: Game and Rule Interpretations” ( $N = 27$ ).
- The children who participated came from several Day Care Centres and Kindergartens in Victoria, B.C.

## PERSPECTIVES

### History of Childhood as History of Education: A Review of Approaches and Sources

R. L. Schnell

*The University of Calgary*

Childhood and family history are developing specialties that promise to integrate many of the latest innovations in contemporary scholarship (Kren, 1977). The history of family life is associated with both the new quantitative history with its use of computers, sophisticated statistical programs, census and other demographic data, and family reconstitution techniques and the new social history's use of social scientific theories and concepts to organize historical data into explanatory propositions. Although some scholars have gone to humanistic, psychological, and particularly psychoanalytic literature for insight into and explanations of human actions, most scientifically minded historians have turned to the behavioural sciences for aid in formulating scientifically testable hypotheses.<sup>1</sup>

While historians of family life are attracted to quantitative and social scientific historical inquiry, historians of childhood have sought their conceptual framework in psychoanalysis or critical history. Childhood historians, in moving away from public events and large scale social data have been forced to consider anew some fundamental questions of the historical enterprise, particularly the relationship between private experience and public events and the use of concepts in the selection and organization of evidence.<sup>2</sup>

Since family life and childhood are associated biologically and socially, many studies span both specialties; however, even these tend to display either a social or a psychological emphasis depending on the primary focus of attention.

As in the case of the "new" history of education, historians have been attracted by the intersection of significant elements of human experience which offer opportunities for understanding institutions, processes, and actions. In this case, however, historians of education share in the pleasant prospect that a new historical specialty will not only enrich their work but that without a proper

consideration of education the history of family and childhood will be seriously truncated.

Educational historians have begun to include materials from the new areas in their course, and in a number of universities, courses clearly identified as family and childhood history have been developed as history of education offerings. With so much activity in North America and Europe, such courses will soon become basic elements in any full-fledged history of education program in Canadian universities. Also the centrality of the areas to so many aspects of human experience suggests that a variety of academic units, e.g., departments of history and literature and faculties of social welfare, will soon be offering their versions of history of childhood and family life.

It seems axiomatic that subjects such as "childhood" and "family" would have a *prima facie* case for inclusion in the professional education of teachers. Whether their focus is social or psychological, such courses provide intending-teachers with a new way of understanding the development of educational structures, processes, and theories. Either as a form of grass roots history of education or as a conceptually organized interpretation of the interactions of private experiences and public events, these courses provide another perspective that promises to link education, and particularly schooling, to deeply felt common human experiences.

Such emphases also promise to humanize the formalistic character of both the older educational historian's preoccupation with the internal logic of institutional pedagogy and the new social historian's obsession with quantification and social scientific concepts. While students are attracted to courses that treat personally significant social categories, the stress on concrete individual experience and the necessity for reconstructing past sensitivities and sensibilities promise opportunities for realizing a major goal of historical inquiry, namely that each era and society must be understood in its own terms. To recognize and to act on the principle that our values and categories are not universally held or natural is a significant result of historical study.

What follows is an explanation of the problems and opportunities facing the educational historian who attempts to incorporate history of childhood and family life into the history of education offerings in faculties of education. For brevity, the term "history of childhood" will include elements of "history of family life" that are relevant to the history of education.

## I

Although it is only the recent interest in childhood and family life that has produced the necessary monographs and articles for university courses, there exists a scholarly literature in closely related areas, e.g., religious and ethnic studies, which have provided a partial basis for the new work and have, in turn, been revitalized by the new specialties. The sudden surge in activity is connected to certain methodological problems in historical scholarship and to the development of strategies to outflank these obstacles.

A major problem is the very nature of history, namely that it is preeminently the study of documents and, in particular, documents related to large historical institutions. Since it is just those institutions that produce and conserve documents, historians have been driven to concentrate on them. Perhaps "driven" is too strong a word because historians have generally preferred public events to



private experiences not only because of practical difficulties but because they have deemed public events to be more important, to be the real point of history, while private experiences are seen as the subject matter of autobiography and biography, literature, psychology, and aesthetics.

Even those historians concerned with plotting the *public* terrain of childhood are faced with problems of suitable evidence. Documents of public events are usually marked by continuity and at least minimal relatedness. Documents of childhood and family life are plagued by discontinuity and uniqueness. Children and families, except for the selfconscious and the powerful are counted among the mute of history, that is, those faceless throngs who provided the energy if not the wittiness of society.

There is another inherent characteristic of history that must be noted. Although many historians talk of the "creative reconstruction of past events," the historical enterprise is more properly and appropriately concerned with the construction of historical knowledge based on a disciplined examination of documents (Meiland, 1965). This observation is based on two considerations. First, it cannot be that "reconstruction of past events" requires the restoration of a particular event in its fullness. Such a restoration could no more be grasped by the historian than by the intelligent contemporary viewer unless, of course, one postulates an historical mind that is divine or at least angelic. Secondly, however, there remains a more fundamental objection to the reconstructionist position. Historians are concerned with providing interpretations of events that link discrete facts into a significant and meaningful narration. A crucial part of the historical enterprise is to determine which facts of the past are historical facts. Decisions in such matters do not arise out of the phenomena but rather are based on assumptions brought by the historian to his work (Leff, 1971).

Two major breakthroughs have radically changed the possibility of history of childhood. One is the development of demographic techniques and statistical methods which allow for the virtual reconstruction of past communities and for the establishment of empirically verified generalizations on such matters as fertility, mortality rates, family and household size, and the distribution of inheritances.<sup>3</sup> This research has called into question many of our beliefs about early modern family life. These beliefs must now be seen as part of the Edwardian mythology surrounding Victorian England and by extension the rest of Europe and North America.

The second breakthrough is perhaps best illustrated by Philippe Ariès' *Centuries of Childhood*. Although historians have been aware that documents include more than the written word and although art historians and classicists have made brilliant use of iconography, architecture, literature and language, to understand classical antiquity and the function of central concepts in Western life, historians studying private but publicly significant experiences have been slow to make use of such evidence and techniques. Ariès demonstrated the effectiveness of the procedures in making manifest changing perceptions of children and the growing sentimentality of a child-centred family and family-centred society. Although Ariès' interpretation has been vigorously challenged, the enduring result of his work was to demonstrate how indirect evidence — given the appropriate conceptual framework — can add a new dimension to historical knowledge by providing the mute of history with the means of speaking to us.<sup>4</sup>

Ariès' approach offers a means of circumventing the traditional problems of the

fragmentary nature of memoirs, autobiographies, and diaries, the unrepresentativeness of materials from the self-conscious portions of élites, and the fact that nearly all accounts of childhood are filtered in one way or another through adult eyes.

With all that, we are still left with the problem of understanding the nature of the relationship between childrearing literature and normal practices of the times. The popularity of such manuals as Dr. Spock's *Baby and Child Care* raises a variety of questions. First, although it is sometimes possible to determine the sales and even circulation rates of these manuals, it is unclear whether they are charting new directions in child care or primarily restating conventional wisdom. Secondly, the existence of child care practices conforming to the advice of such manuals do not imply unambiguous connections — even in the cases of readers of the manuals — for we may interpret such correspondence as the reflection of traditional patterns given a new guise by writers of child care literature. Thirdly, practices that differ from traditional ones may have arisen from conditions quite unrelated to the manuals. Finally, popular child care manuals require careful examination in light of the consideration that there is more to them than recommendations for child care. That is, their popularity needs to be understood in connection with a whole series of social interests, needs, and relationships.

It is in the social analysis of childrearing and etiquette literature that Ariès has again demonstrated the possibility of getting inside the mentality of various social strata over a long historical period by determining the principle categories that have shaped men's perception of the world.<sup>5</sup> Ariès has reemphasized the significance of concepts in framing not only our social world but also our historical interpretations and the centrality of the construction of historical knowledge in the historical enterprise.

## II

A remarkable trait of history of childhood is the lack of dogmatism generally and the diversity of approaches in course organization and research. This diversity is an expression of two fundamental historical principles. First, since history is the interpretation of documents, the wide range of documents has created a corresponding array of approaches and interpretations. Secondly, historians have frequently sought for the representative case, not in the sense of the average but rather as the unique and even eccentric embodiment of certain social characteristics. Thus, focusing on the exceptional enables the historian to put in bold-relief the peculiarities of a given historical age.

With these principles in mind, we can more fully appreciate the possible and actual methods of dealing with the problems of childhood history. One approach that connects history of childhood to the traditional study of pedagogues and their theories calls for an examination of the literature in the light of the concept of childhood. Unlike the older practice of analyzing educational theories in terms of standard pedagogical interests, the use of significant social concepts provides a link between the theorists and the ideologies of their times. It also liberates us from the ahistorical practice of reading our cultural concerns and understandings into the past. Although presentism is probably too pervasive to be totally excluded from our work, the history of childhood should alert us to the dangers of assuming the existence of familiar institutions and perceptions. In this instance, we are able to strip away the sentimentality that surrounds discussions of children and family life



and to avoid the mistake of using an idealized “modern” image of the family and family life by which to judge all historical cases.

An examination of the pedagogues allows us to estimate the extent to which such central modern concepts as childhood, family, and adolescence were significant elements in the thinking of a particular historical era. There is also considerable opportunity for re-analyzing pedagogical concerns in terms of these social categories and thereby re-connecting what often appears to be abstract and lifeless to institutional and intellectual considerations, i.e., the very things of which human experience is composed.<sup>6</sup>

A second major avenue into childhood history is through literary sources — an approach based on two assumptions. First, cultural historians are interested in the “representative” man or woman of a particular time and place. Although few would argue that fictional and poetic sources accurately depict historical situations, they do provide us with two kinds of understanding: (1) consciously or not, literary writers offer a way of looking at their world from within, i.e., the perspective of those who are a part of that world; and (2) since literature can only reflect, examine, or criticize a part of that world, then the authors must concentrate their consciousness on selected portions of their society. In these ways, writers reflect society and human nature through a literary prism to separate out and thereby to allow us to see uniquely, if distortedly, the pieces of a living world. Literature gives us a means of seeing not the commonplace but those characteristics of a society’s life that give it a style and tone, i.e., that are representative of it.

Secondly, literary materials are frequently produced by “representatives” of not only their own time but of the future, i.e., those who anticipate attitudes and sensitivities that are only an embryonic form in society. Although not all sensitivities come to define or dominate the thought and practice of the future, they do show us how certain forecasts are born, die, or come into organic relationship with enduring institutional and intellectual patterns.

Besides the original sources for such inquiry, there are major studies that illustrate significant developments in literary thought about children, childhood, and sensibility. In particular, studies of literature about and for children document not only patterns of social perceptions and expectations but also the ways in which literary and intellectual perspectives work their way into ordinary language and actions.<sup>7</sup>

A third approach to childhood history draws heavily upon the theories and concepts of psychology and particularly psychoanalysis. There is a distinction between the use of psychological literature as with literary materials in order to understand a given time and place and the use of psychological theories and concepts as the means of organizing and interpreting historical phenomena. In the first instance, psychology is problematic, that is, another product of human imagination; while in the second, it provides directive assumptions that guide the selection of data and their interpretation.

The most common basis for a psycho-history of childhood is psychoanalysis. Frequently the theoretical framework comes not directly from Freud but rather as modified by Erik H. Erikson, who has postulated a theory of psycho-social development marked by eight stages and specific crises. Erikson has also argued for an essentially benign relationship between the human organism and its social environment. The clarity of the Eriksonian schema with its accommodation of



cultural and historical conditions has attracted the attention of historians of childhood.<sup>8</sup>

The use of a psychological framework brings with it certain difficulties for historians who are required to accept certain assumptions about human nature and society that might otherwise be considered objects of inquiry. Also, given the multitude of human experiences and the fact that psychologies are not created outside those experiences, historians are able to provide data that illustrate the "truth" of the psychologies. In the case of a truly universal science, e.g., psychoanalysis, that claims to explain all human experience, one cannot conceive of historical evidence that would count against it. Consequently, much of psycho-history is ultimately the use of historical data to illustrate psychological theories. The principle attraction of the psycho-historical approach is its promise to give significance to scattered and idiosyncratic recorded experiences; however, if psychologies are to open new avenues of historical insight and inquiry, they must be used to suggest new points of historical research and analysis and not as monuments to which historical gingerbread is added for aesthetic reasons.

A recent example of the application of psychoanalysis in a universal theory of historical development is the work of Lloyd deMause and his associates in the Association for Psychohistory. Through the *Journal of Psychohistory* (formerly the *History of Childhood Quarterly*), sponsored publications, and conferences, the deMauseans have promoted a psychogenic theory of childhood that is also a psychogenic theory of history. The psychogenic theory has the twin advantages of a systematic psychoanalytic explanation for changing patterns of child care and institutional development and a stress on human relationships (and especially family ones) that so dominates the mind and popular social science of North America. The familiar themes of battered and abused children and wives, the lack of empathy, and the widespread use of guilt and psychological manipulation are some of the links between psychogenic theory and the educated mind of the West. As an example of Whig history with its emphasis on the victory of modern attitudes and practices over the primitive, psychogenic theory is concerned with discerning the roots of the present in the past. Consequently, psychogenic theory postulates a social equivalent of biological evolution that is a mixture of inevitability and moral progress.<sup>9</sup>

A fourth approach to childhood history is found in the studies of the institutionalization of children, adolescents, and deviants. In large part, these studies are a response to major problems in contemporary society. Confronted by contradictions in social policy and practice, historians have sought their sources in the movements for child-saving in the nineteenth and early twentieth centuries. These scholars have demonstrated a massive transformation of western society in terms of the creation of institutions that came to control more and more of the population and the intellectual and social conditions that made such institutionalization reasonable and desirable. Although they have provided us with exciting and imaginative examinations of these changes, these historians have not been able to transcend the contradictions that are part of the contemporary thinking about related social issues. Also, even if the massive transformation of the past two centuries is conceded, we still require an organizing principle that will conceptually link childhood before the transformation with the contemporary status of children and adolescents.<sup>10</sup>

A final approach to childhood history involves the use of "case studies" in

depicting the coming-of-age in a variety of urban and rural surroundings. Emphasizing the shock of recognition and the possibility of a recoverable past, the literature has moved in three streams. First, since the early 1950s there has been a strong interest in historical urban and regional studies. Some of these have incorporated the full array of quantitative historical methodologies, while others have relied on traditional historical techniques for recreating a past milieu. A second avenue to past communities is the considerable body of older sociological studies dating from the 1920s. Although most of the work is American and British, there are sufficient Canadian and European studies to allow for the cross-cultural comparisons. A third variety of case studies is the autobiographical and journalistic recreations of the recent past.<sup>11</sup>

There are difficulties with all three approaches. First, the use of such studies can easily degenerate into a rag bag of odds and ends strung together by time. Second, there are again the problems of representativeness, fragmentation, distortion by subsequent experience, and the necessity of connecting private experience and public events. Finally, the last two forms of case studies are usually static slices of life without dynamics to account for changes over time.

### III

All approaches to childhood history have been faced with the same historical and pedagogical question, namely, How can disparate data be organized into a meaningful and significant interpretation?

An increasingly popular integrating principle has been psychological models of human development and especially those postulated by Erik H. Erikson and Lloyd DeMause. As mentioned earlier, the rigorous application of psychological models to historical data introduces an element of determinism into historical interpretation that most historians find repugnant and raises the possibility that the activity underway is essentially psychology rather than history.

Some historians are attempting to establish significant patterns of human life, e.g., sentiment, morality, and mass behaviour, by a close examination of a wide range of historical studies (Shorter, 1975). Even with the determination of these patterns, one is still left with the necessity of organizing the empirical generalizations into an historical interpretation that has coherence, clarity, and point. What these scholars have done is to produce historical evidence that is more abstract and generalized than most historical facts. What they have not done is to provide an interpretation that uses the generalizations.

A third — and pedagogically more useful — approach is best reflected in the work of Philippe Ariès who has made extensive and subtle use of “concepts” in ordering an impressive variety of historical evidence. In a sense, the “conceptual framework” approach borrows more from the recent revolution in philosophy than from social science in that the thrust is to establish the essential characteristics of major social categories, e.g., childhood, family, and adolescence, and then to explore the historical development of the concepts into their contemporary form. By focusing on central concepts in history of childhood, the conceptual framework approach provides an integrative principle for case studies and evidence drawn from pedagogical thought and literary and intellectual sources.<sup>12</sup>

This approach rests on several assumptions. First, that the end of the historical enterprise is not the reconstruction of the past but the construction of historical



knowledge. Documents, which are the evidential base for historical knowledge, include not only written records but also the monuments of the past, e.g., architecture and iconography. Secondly, that historians are concerned with historical interpretation rather than causal explanations. As a result historians are interested in constructing historical interpretations that are dependent on historical documents for evidence but that are essentially creative activities controlled by the dominant views of human nature and society and by the canons of logic. Thus, we are able to account for widely differing interpretations of historical phenomena — even where similar evidence is used — by understanding the assumptions underlying each interpretation.

Thirdly, underlying this view of the historical enterprise is the more basic assumption that the human mind is “creative and self-determining, playing a constructive role in life and experience rather than passively responding in mechanical fashion, to the promptings of external or internal stimuli” (Gardiner, 1974) and that the way in which the mind plays such a “role” is by means of concepts and especially social categories that give order and point to experience. Thus, to grasp what is entailed in key concepts over time is to understand the point of institutions and human activities, that is, to understand what they are intended to do. It should be noted that the conceptual framework approach is based both on the centrality of concepts in human experience and in the construction of historical interpretations of that experience. In that sense, the approach and the theory are consistent.

Fourthly, that the criteria for an acceptable or adequate historical interpretation are that the interpretation “must accord with the evidence . . . must show the connection between the events it interprets or at least their significance in relation to their context . . . and . . . , should not presuppose unlikely, implausible or unacceptable assumptions” (Leff, 1971). The crucial element is the matter of assumptions for not only do different assumptions lead to different conclusions from the same data but assumptions often direct us to select different evidence. The conceptual framework approach requires us to be critically self-aware in understanding and explaining the process of historical inquiry *and* the process by which we construct our words.

Fifthly, that a good history course is essentially an extended historical interpretation. As a result, the central characteristic of the course must be a critical use of concepts, logic, and evidence in light of the nature and limitations of any historical interpretation. Finally, that a history of childhood course as an extended historical interpretation is conducive to the development of skills of historical analysis and synthesis. Such skills allow us to see our world differently in ways that we most likely would not have and to confront that world with all its risks.

#### IV

It is now necessary to outline briefly the organization of a history of childhood course using the conceptual framework approach. Given the importance of assumptions, the approach requires the examination of the dominant models of childhood in order to explicate the assumptions, logic, and evidence used. There are available three major interpretations that allow for the full use of conceptual analysis, that is, the psychogenic theory of Lloyd deMause, the psycho-social developmentalism of Erik H. Erikson, and the cultural relativism of Philippe Ariès. It is not possible here to offer any explication of the three models. Rather what



follows is a brief description of the use of the models in an undergraduate course in a faculty of education, a course which is organized so as to focus specifically on the relationship of childhood, family, and institutional education.

After some introduction to the nature of the historical enterprise, students are initiated into an examination of the three models by a combination of assigned readings and lectures that provide a systematic explication of each model in turn. After the deMausean and Eriksonian models have been introduced it is possible to exploit them pedagogically by comparing them in terms of evidence used and the interpretations given for similar data. This allows students to comprehend the importance of assumptions and to begin the first steps in historical analysis by examining the logical and practical consequences of different assumptions. Extracts from primary and secondary sources for class analysis can also be employed to demonstrate the process of determining assumptions, use of evidence, and argument.

We then move on to Ariès' *Centuries of Childhood* which requires detailed classroom analysis in order to realize its full potential. In particular, it is necessary to engage in an analysis of Ariès' concept of childhood by stripping away all but its essential meaning. With the analysis completed, the class can now read Ariès with increased understanding with respect to (1) the connection between childhood and family, (2) the relationship of schooling to childhood and family, and (3) the functions of religious, social, and intellectual movements in shaping institutions and processes serving children and youth. Slides of the iconographic evidences used by Ariès and critical articles examining various parts of his interpretation are used to demonstrate the varieties of historical evidence and their limitations as well as the ways in which different assumptions influence historical judgments.<sup>13</sup> It is then possible to engage in some comparative analysis of the three models of childhood by examining the ways in which the same historical phenomena, e.g., swaddling and apprenticeship, are explained. Once again extracts of primary and secondary sources are a convenient means of conducting in class analysis of historical propositions.

In describing the conceptual framework approach, we have emphasized the triads of concepts, logic, and evidence in analyzing models of childhood and historical statements. We should not, however, lose sight of the final end of this initial work, namely, the construction of historical knowledge in the form of an interpretation. An interpretation is basically an act of synthesis — the organization of a logically coherent argument based on relevant evidence and reasonable assumptions. One of the assumptions underlying the course was that it was to be an interpretation that linked together childhood, family, and schooling. The examination of the works of deMause, Erikson, and Ariès provided the means to connect the three institutions prior to the establishment of compulsory schooling in the nineteenth and twentieth centuries.

The course is an interpretation of the relationship of childhood, family, and common schooling which argues that changes in childhood and family life since the Middle Ages have resulted in removing children and youth from the rigours of adult society. This removal increasingly limited the liberty of children and youth and asserted that their natural condition ought to be one of dependence, protection, segregation, and delayed responsibility. In order to bring about such a change in the status of children and youth, more than the transformation of the family from an economic to an affectionate institution was needed. What was

required was a universal institution that had as its major responsibility the rearing of children and youth in a protected and controlled environment. The obvious agency for such control was the school and the means for institutionalizing all children was the common school.

In order to complete our synthesis, we must consider the major interpretations of the use of common schooling. After an examination of the four major positions — Whig, social control, conflict, and anxiety<sup>14</sup> — in which their principal characteristics were established, we then turn to the connection between the rescue motive developed in the analysis of the models of childhood and the interpretations of common schooling.

The course has now reached its conclusion with the final argument that the common school was the most ordered of rescues by providing a compulsory childhood for all children. By focusing on the essential criteria of the concept of childhood, we have been able to develop the links among three institutions — childhood, family, and school — that allow us to offer an historical interpretation of the rise of compulsory institutionalization of children and youth. In so doing, the course itself is a cohesive unit that subjects various historical interpretations to the rigours of critical analysis and introduces newcomers to historical methodology and to the major problems inherent in the discipline itself, e.g., matters of cause, explanation, evidence, and assumptions. Likewise this analysis proves to be the means for a compatible examination of the popular interpretations of the rise of the common school.<sup>15</sup> Thus, the methods of analysis and synthesis have provided us with the means of linking together three significant institutions of modern society — childhood, the family, and the school.

#### Notes

1. Shorter (1975) has provided the best general introduction to the quantitative study of family life. The leading periodicals are the *Journal of Interdisciplinary History* and the *Journal of Family History*.
2. In addition to the seminal work of Ariès (1962), we now have major studies of children in Canada (Sutherland, 1976), England (Pinchbeck & Hewitt, 1971-1973), and the United States (Cable, 1975), as well as youth in England and Germany (Gillis, 1974) and America (Kett, 1977).
3. A general survey of historical demography is given by Hollingsworth (1969).
4. An excellent example of this scholarship is Jaeger (1943-1945).
5. The etiquette and courtesy literature has been critically examined in Barker (1967), Mason (1971), Ferguson (1960) and Esler (1966). A significant aspect of this literature and its nineteenth century equivalents (cf. Crandall, 1969) is that they served as a means of re-socializing young men moving from one social stratum to another as well as introducing knowledge of new social roles into schools and universities.
6. Although major educational theories and practitioners of the early modern Europe (1650-1850) anticipated many of the changes in sensibilities and practices in child care, there are no modern studies examining Comenius, De la Salle, Francke, Pestalozzi, Herbart, and Froebel from either the perspective of institutionalization or the growing concern for psychological control of children and adolescence.
7. The central work is by Coveney (1967); however, one should not neglect Walsh (1966) nor useful studies of children's literature by Avery (1965), and Kiefer (1948).
8. The Eriksonian study of greatest interest to the historian of childhood is *Childhood and Society*. For assistance in understanding Eriksonian theory and its implications, consult the study by Roazen (1976).
9. L. deMause has used the *Journal of Psychohistory* (formerly the *History of Childhood Quarterly*) to expound his view on psychohistory (1975a, 1977) and his psychogenic theory of childhood (1974, 1976). He has also edited an important collection of original historical studies (1975b).
10. Exemplary studies of institutionalization include Rothman (1971), Mennel (1973), Platt (1977), and Schlossman (1977).
11. The classic sociological community studies for the United States and Great Britain are conveniently surveyed by Stein (1964) and Frankenberg (1966). One might wish to supplement these with such community studies as Wylie (1964) on France and Burnet (1951) on Alberta.

Although in North America most attention has been on colonial New England (Demos, 1970), historians have studied such diverse communities as nineteenth century Hamilton, Ontario (Katz, 1976) and frontier Wisconsin (Curti, 1959).



The popular accounts range from descriptions of village (Blythe, 1970) and slum (Roberts, 1973) life to the reconstruction of childhood experiences (Wheelis, 1958).

12. Although there is considerable interest in categorization among some sociologists of knowledge, the analysis employed here is more closely related to analytic philosophy (Peters, 1966) than to sociology.
13. Besides *Centuries of Childhood*, Ariès has expounded his views on the mid-twentieth century family (1975). Also see Mousseau's (1975) interview of Ariès. Useful critiques of his interpretation are found in Forsyth (1976), Hanawalt (1977), and Hunt (1970).
- Materials for analyzing and extending Ariès interpretation are found in Cleverley & Phillips (1976), and Wishy (1968).
14. Examples of the four interpretations abound in the literature. The most notorious North American example of educational Whiggery is Cubberley (1919). Pedagogically Katz (1975) is the best introduction to the assumptions and logic of the social control theory. Excellent examples of the conflict interpretation are provided by Simon (1960-1974) and Laqueur (1976). At present the anxiety model has been explicated at length only by Finkelstein (1975, 1976).
15. For a detailed explication of this interpretation, see Schnell (1979).

### References

- Ariès, P. *Centuries of childhood: A social history of family life*. New York: Alfred A. Knopf, 1962.
- Ariès, P. La famille: A report from France. *Encounter*, 1975, 45, 7-12.
- Avery, G. *Nineteenth century children — Heroes and heroines in English children's stories, 1780-1900*. London: Hodder and Stoughton, 1965.
- Barker, E. *Traditions of civility: Eight essays*. Hamden, Conn.: Archon Books, 1967.
- Blythe, R. *Akenfield: Portrait of an English village*. New York: Dell, 1970.
- Burnet, J. *Next year country: A study of rural social organization in Alberta*. Toronto: University of Toronto Press, 1951.
- Cable, M. *The little darlings: A history of child rearing in America*. New York: Charles Scribner's Sons, 1975.
- Cleverly, J., & Phillips, D. C. *From Locke to Spock: Influential models of the child in modern western thought*. Carlton: Melbourne University Press, 1976.
- Coveney, P. *The image of childhood*. Harmondsworth: Penguin Books, 1967.
- Crandall, J. C. Patriotism and humanitarian reform in children's literature, 1825-1860. *American Quarterly*, 1969, 21, 3-22.
- Cubberley, E. P. *Public education in the United States*. Boston: Houghton Mifflin, 1919.
- Curti, M. *The making of an American community: A case study of democracy in a frontier county*. Stanford: Stanford University Press, 1959.
- deMause, L. The evolution of childhood. *History of Childhood Quarterly*, 1974, 1, 503-606.
- deMause, L. Independence of psychohistory. *History of Childhood Quarterly*, 1975, 3, 163-200. (a)
- deMause, L. (Ed.). *The history of childhood*. New York: Harper Torchbooks, 1975. (b)
- deMause, L. The formation of the American personality through psychospeciation. *Journal of Psychohistory*, 1976, 4, 1-30.
- deMause, L. The psychogenic theory of history. *Journal of Psychohistory*, 1977, 4, 253-267.
- Demos, J. *A little commonwealth: Family life in Plymouth Colony*. New York: Oxford University Press, 1970.
- Erikson, E. H. *Childhood and society* (2nd ed.). New York: W. W. Norton, 1963.
- Esler, A. *The aspiring mind of the Elizabethan younger generation*. Durham: Duke University Press, 1966.
- Ferguson, A. B. *Indian summer of English chivalry*. Durham: Duke University Press, 1960.
- Finkelstein, B. Pedagogy as intrusion: Teaching values in popular primary schools in nineteenth-century America. *History of Childhood Quarterly*, 1975, 2, 349-378.
- Finkelstein, B. Fear of childhood: Relationships between parents and teachers in popular primary schools in the nineteenth century. *History of Childhood Quarterly*, 1976, 3, 321-335.
- Forsyth, I. H. Children in early medieval art: Ninth through twelfth centuries. *Journal of Psychohistory*, 1976, 4, 31-70.



- Frankenberg, R. *Communities in Britain: Social life in town and country*. Harmondsworth: Penguin Books, 1966.
- Gardiner, P. (Ed.). *The philosophy of history*. London: Oxford University Press, 1974.
- Gillis, J. R. *Youth and history*. New York: Academic Press, 1974.
- Hanawalt, B. H. Childrearing among the lower classes of late medieval England. *Journal of Interdisciplinary History*, 1977, 8, 1-22.
- Hollingsworth, T. H. *Historical demography*. Ithaca: Cornell University Press, 1969.
- Hunt, D. *Parents and children in history*. New York: Basic Books, 1970.
- Jaeger, W. *Paideia: The ideals of Greek culture* (3 vols.). Oxford: Basil Blackwell, 1943-1945.
- Katz, M. B. *Class, bureaucracy, and schools* (rev. ed.). New York: Praeger, 1975.
- Katz, M. B. *The people of Hamilton, Canada West: Family and class in a mid-nineteenth-century city*. Cambridge: Harvard University Press, 1976.
- Kett, J. F. *Rites of passage: Adolescence in America, 1790 to the present*. New York: Basic Books, 1977.
- Kiefer, M. *American children through their books*. Philadelphia: University of Pennsylvania Press, 1948.
- Kren, G. M. Psychohistory in the university. *Journal of Psychohistory*, 1977, 4, 339-350.
- Laqueur, T. W. *Religion and respectability: Sunday schools and working class culture, 1780-1850*. New Haven: Yale University Press, 1976.
- Leff, G. *History and social theory*. Garden City: Anchor Books, 1971.
- Mason, J. E. *Gentlefolk in the making: Studies in the history of English courtesy literature and related topics from 1531 to 1774*. New York: Octagon Books, 1971.
- Meiland, J. W. *Scepticism and historical knowledge*. New York: Random House, 1965.
- Mousseau, J. The family, prison of love. *Psychology Today*, 1975, 53-58.
- Peters, R. S. *Education and ethics*. London: George Allen & Unwin, 1966.
- Pinchbeck, I. & Hewitt, M. *Children in English society* (2 vols.). Toronto: University of Toronto Press, 1971-1973.
- Platt, A. *The child savers: The invention of delinquency* (rev. ed.). Chicago: University of Chicago Press, 1977.
- Roazen, P. *Erik H. Erikson: The power and limits of a vision*. New York: Free Press, 1976.
- Roberts, R. *The classic slum: Salford life in the first quarter of the century*. Harmondsworth: Penguin Books, 1973.
- Rothman, D. J. *The discovery of the asylum: Social order and disorder in the new republic*. Boston: Little, Brown, 1971.
- Schlossman, S. L. *Love and the American delinquent: The theory and practice of "progressive" juvenile justice, 1825-1920*. Chicago: University of Chicago Press, 1977.
- Schnell, R. L. Childhood as ideology: A reinterpretation of the common school. *British Journal of Educational Studies*, 1979, 27, 7-28.
- Shorter, E. *The making of the modern family*. New York: Basic Books, 1975.
- Simon, B. *Studies in the history of education* (3 vols). London: Laurence and Wishart, 1960-1974.
- Stein, M. R. *The eclipse of community: An interpretation of American studies*. New York: Harper Torchbooks, 1964.
- Sutherland, N. *Children in English-Canadian society: Framing the nineteenth-century consensus*. Toronto: University of Toronto Press, 1976.
- Walsh, W. *The uses of imagination: Educational thought and the literary mind*. Harmondsworth: Penguin Books, 1966.
- Wheelis, A. *The quest for identity*. New York: W. W. Norton, 1958.
- Wishy, B. *The child and the republic*. Philadelphia: University of Pennsylvania Press, 1968.
- Wylie, L. *Village in the Vaucluse* (rev. ed.). New York: Harper Colophon Books, 1964.
- Zuckerman, M. *Peaceable kingdoms: New England towns in the eighteenth century*. New York: Alfred A. Knopf, 1970.

## BOOK REVIEWS

### *Rejoinder*

*Rejoinder to Professor L. K. Schubert's review of John H. Andreae, Thinking with the Teachable Machine. London, England: Academic Press Inc., 1977, 178 pp. The Alberta Journal of Educational Research, Vol. XXIV, No. 4, December, 1978, 291-293.*

Authors must have a thick skin and critics will always occupy a privileged position when they slate the work of others on the basis of a casual understanding. In an artificial intelligence (AI) journal, L. K. Schubert's review of my book, "Thinking with the Teachable Machine" (1977, Academic Press) might have been seen as the powerful AI orthodoxy rejecting an heretic, but in the December issue of *this Journal*, his words may well have puzzled readers with educational interests.

The development of my learning system, PURR-PUSS, did not proceed from the orthodox assumption in AI that "learning is a second-order effect" (A. Newell and H. A. Simon: Human Problem Solving, page 7, Prentice-Hall, 1972), but from the opposite assumption that learning comes first and provides a substrate for intelligence. Those who can entertain this second assumption will appreciate the modest efforts that have resulted in PURR-PUSS and will be surprised to find no astounding claims in my book. They will find that Schubert's review is full of inaccuracies that miss the point. However, there is one criticism which is so strong and so wrong that I cannot allow it to go unanswered. Were he right, the system would be useless.

In making the following assertion, Schubert is denying the inherent mode of operation of PURR-PUSS: "This illustrates a fundamental weakness in Andreae's prediction scheme: its inability to *generalize* over sequences and their relationships, i.e., its inability to form concepts."

PURR-PUSS interacts with the real or simulated world via several channels of input and output, which allow "her" to sense a variety of modes of input (e.g., sight, sound, and touch) and to effect a variety of modes of output (e.g., to eyes, voice, and limbs). The better the robot body which can be provided, the more useful will be the input and output channels.

For each channel, a fixed number of the most recent events on the channel comprise the "context" for a fixed-length predictor. The collection of contexts, called the "multiple context," is PURR-PUSS's information about the current situation. Her decisions to act are derived by simple majority-evidence logic from the several predictions of the predictors. Since as few as two predictions need agree for an action to be selected, much of the time PURR-PUSS is deciding with only part of her multiple context. Her decision mechanism *generalizes* from the partial



situation, represented by a few of the contexts, to the current situation, represented by the whole multiple context. Once the action has been taken, it is associated in memory with the whole multiple context and, if that situation occurs again *in toto*, all the contexts will *discriminate* in favour of the same action.

Schubert does not seem to have understood this process, explained in detail in chapter 4, or he would not have attributed the limitations of a single predictor to the multiple context.

The approach to intelligence via learning is slower and less spectacular than the programming of cleverness. Learning systems have a very long way to go. My book was written to encourage others to take this approach with a teachable machine interacting in the real world. Orthodox AI may continue to accumulate expertise about “specific mechanisms for solving problems, proving theorems, understanding language, manipulating blocks, and so forth,” as Schubert recommends, but my intuition is that they will not construct an intelligent machine that way. In my view, learning will be as vital to a really intelligent machine as growing is to a flower. Artificial flowers have many advantages and may well compete with nature for colour, texture, and fragrance, but they remain dead. The products of orthodox AI are magnificent and in many ways superior to the human brain, but they lack learning, the double-helix of intelligence.

John H. Andrae

*Department of Electrical Engineering  
University of Canterbury  
Christchurch, New Zealand*

A THEORY OF EDUCATION. *By Joseph D. Novak.* Ithaca, New York: Cornell University Press, 1977, 291 pp.

A theory of education is an elusive ideal. Part of the elusiveness is caused by the many meanings of the word “theory.” At one extreme theories, particularly in the natural sciences, are highly formalized and abstract sets of statements that help to explain and predict phenomena. The kinetic theory of gases and Bohr’s theory of the atom are examples of this type of theory. At a somewhat broader level we find general areas of research labelled theories as in physical or nuclear theory. Sometimes “theory” refers to general patterns of explanation which tell us what the main concerns in an area are but which do little by way of providing specific grounds for making predictions as in functional theory in sociology, or Piagetian theory or Gestalt theory in psychology. Still another sense of “theory” is that of a general abstract program which is claimed should be adopted, as in Marxist theory. Finally at the opposite extreme from specific scientific theory is the use of the term to identify hypotheses, or even intuitive hunches, about the world which are often unsupported as in the claim that a person has a theory of how government will react to declining school enrollments. The other part of the elusiveness of a theory of education is caused by the subject matter of education. Much of our knowledge about education is empirical or factual. Research into questions such as whether homogeneous grouping of students increases achievement levels or whether higher order questioning by teachers improves understanding can be done by purely empirical techniques. But the empirical aspect of education does not exhaust the concept. Education is a value laden notion. One cannot engage in activities designed to educate others without making the judgment that what is to be learned is worthwhile. Questions about improving achievement levels or understanding are only meaningful *educational* research questions because it has been judged that



these are desirable goals. So a theory of education must do justice to the dual aspect of its subject matter and must make clear in what sense it is a theory. It is little wonder, then, that the whole notion of a theory of education is one wrapped in controversy and which presents an elusive ideal to researchers.

It is into this treacherous area that Novak steps with confidence, insight and a wealth of experience. His plan is to present a philosophical position, a theory of learning, and a theory of curriculum which together comprise a theory of education. He presents some implications of this theory for the planning of instruction and reports empirical studies which confirm the theory. My general impression of the book is that the positions on which his theory are based are plausible, helpful, and probably right, but that they do not add up to a theory of education. That is, the ideal remains as elusive as ever.

Before commenting on the argument as a whole let me briefly recount the constituent parts of the theory. The philosophic theory is that of Stephen Toulmin in *Human Understanding*. In the one volume published of the three planned, Toulmin is concerned with the historical evolution of human knowledge and understanding, i.e., the growth of concepts. The motto of the book, taken from Kierkegaard, nicely expresses the theme of Toulmin's work: "Concepts, like individuals, have their histories and are just as incapable of withstanding the ravages of time as are individuals." This is not the place to summarize or evaluate Toulmin's masterful work. Suffice it to say that it puts concepts at the centre of attention and shows that concepts are not immutably fixed, but that they evolve over time just as biological species do.

The psychological theory of learning that is the heart of Novak's theory also centres on concepts. It is David Ausubel's theory of meaningful learning which emphasizes that learning in order to be meaningful must fit into a conceptual pattern. The psychological entities into which new information is assimilated are subsuming concepts in the cognitive structure. It is the existence of such subsumers that makes learning meaningful. For this reason, Ausubel stresses the use of advance organizers in instruction. Again this is not the place to go into the details of Ausubel's work, but it is important to note that Ausubel, like Toulmin, puts the growth of concepts at the centre of his scheme. For Toulmin the concern is with human knowledge and understanding in general. Ausubel deals with an individual's knowledge when he stresses the assimilation of new information within the knowledge structure of the person's cognitive structure.

The third main component is Mauritz Johnson's model of curriculum and instruction. Novak rightly utilizes Johnson's distinction between curriculum and instruction and benefits from the clarity that Johnson has injected into curriculum discussions. Novak also points out that traditional learning theories say little if anything to the planning of curriculum and instruction. He then goes on to do a very nice job of showing the congruence between Ausubel's learning theory and Johnson's model. Another element pointed out at this time is that the discussion and research into the structure of knowledge fit in with the theory. With Ausubel's concern for the individual's cognitive structure it is natural to search out the structure within bodies of knowledge.

So Novak has provided us with an integrated picture of education. It centres on Ausubel's theory of meaningful learning with support from Toulmin's discussion of the growth of concepts and with relation to Johnson's model of curriculum and instruction. Built into this is a great deal of insight on such topics as psychological issues in education, the planning of instruction, and the future of education. All this is supported by the report of the research done by Novak and his co-workers. But I do not think that it all adds up to a theory of education. Let me explain.

The first reason for not having a theory of education is the heavy emphasis on conceptual learning. Education includes at the very least not only conceptual learning, but the learning of skills, facts, theories, attitudes, habits, norms, tastes, and standards of thought, action and decision. So a theory of education must make provision for these. Novak does not limit his discussion to conceptual learning; he does make mention of both skill and affective development, but the main emphasis is on the learning of concepts. To the extent that other types of learning found in education are given short shrift, the proposed theory is weakened.

There is also a sense in which the reliance on conceptual learning creates vagueness. The notion "concept" has a very loose meaning and is used in a variety of ways. Novak defines "concept" as some regularity or relationship within a group of facts and designated by some sign or symbol. Such a definition does not seem to be of much help, for the terms "regularity or relationship" do not seem much clearer than does the term "concept." For example, there is a regularity of sorts in the placement of objects on my desk but I am at a loss to know what concept is exemplified by such a regularity of facts. Also, I have the concept of a unicorn although it does not describe a regularity or relationship within facts. As Toulmin says, "The term 'concept' is in danger [of] becoming an irredeemably vague catch-all."

The second problem that I see is that Novak's theory ignores completely the social aspect of education. Novak is concerned with learning theory, of how an individual learns. This is of course a crucial part of education. But education typically takes place in groups of at least two — a teacher and a pupil — and usually in much larger groups; education always takes place in a society. Novak does not mention these factors. In order to capture all of its manifestations, a theory of education must include a social psychology and a sociology.

My final reason for holding that Novak fails to provide a theory of education stems from the normative aspect of our concept of education. To say that someone is educated is not just to describe some fact about the person, it is also to make the judgment that what the person has learned is valuable or is regarded as valuable. Not everything that can be learned counts as education. And Novak gives us no criteria for distinguishing between learning in general and that learning which is educative. This point also becomes apparent in the discussion of curriculum. His starting point in curriculum is a set of unspecified criteria for selecting knowledge to be taught. But this is precisely the contentious issue in curriculum planning. What knowledge is of most worth? Why do we agree that it is acceptable to teach cooking but not bank robbing? These are value questions and with Novak's eye on a natural science model of theory, there is no room for values in his theory of education. But value questions lie inextricably at the heart of educational matters, and a theory of education must deal with these questions. So, because of Novak's preoccupation with the question of how individuals learn concepts, I feel his proposed theory is too narrow in its concerns to be called a theory of education.

Although I make this criticism I want to stress again that this is a most valuable book which succeeds on many other levels. Two most welcome points I would like to mention. We owe thanks to Novak for introducing Toulmin's work to educational writing. Toulmin's careful and profound analysis of knowledge is a most welcome antidote to the strident and superficial writings on knowledge that are sometimes found under the name of sociology of knowledge. Novak's criticisms of behaviourism are clear, to the point, and well taken. He shows that the behaviourist viewpoint is inadequate for supporting in any way a general theory of education. Although I have criticized Novak for not paying enough attention to the normative aspects of education, by being concerned with *meaningful* learning he



gets much further down that road than do the behaviourists who are concerned with learning.

The book is well and clearly written. Two aspects, though, caught my attention. There is a curious inconsistency in the use of person. He switches between the first and third person pronouns in a jarring way. Also there are lapses in tone that strike the reader. In the midst of a difficult exposition of Ausubel's learning theory there is the sudden claim that Ausubel's monotonous speaking style has not aroused enthusiasm for his ideas. One wonders why this is relevant. Speaking style, typically, has nothing to do with the validity of ideas, and if in educational psychology it is a relevant concern, perhaps we should not look to educational psychology as a basis for a theory of education. The book is printed very well with the virtual absence of typographical mistakes. But there is one which is such a delight that I must mention it. In a discussion of computer assisted instruction (CAI) we find the following claim: "Widespread use of CIA is not likely much before the next century." I suppose that for this at least we should be grateful.

Allen T. Pearson  
*Department of Educational Foundations*  
*The University of Alberta*

RESEARCH IN EDUCATION (3rd Ed.). By J. W. Best. Englewood Cliffs, N.J.: Prentice-Hall, 1977, 403 pp.  
EDUCATIONAL RESEARCH IN CLASSROOMS AND SCHOOLS. By L. Cohen. London: Harper and Row, 1976, 426 pp.

These two volumes, presented as introductory texts in educational research, represent two quite divergent approaches to the topic. *Research in Education* represents the conventional or traditional (in North America, at least) treatment with the stated objective of showing the beginner how to carry out a research project and present it in correct thesis style. In contrast, as the title indicates, the objective of Cohen's book is to introduce the reader to the purposes and techniques of observing pupils and teachers in school and classroom settings. Best mentions his target as the graduate worker in education; Cohen's intended readership is wider, specifically "students reading for B.Ed. degrees and for Diplomas and Higher Degrees in Education."

The disparate content of these two volumes may perhaps best be seen by comparing the 10 chapters of Best with Cohen's 8 sections:

BEST	COHEN
1. The Meaning of Research	1. Observing Pupil Differences
2. Selecting a Problem and Writing a Research Proposal	2. Observing Teacher Differences
3. The Use of Reference Materials	3. Observing Pupils' Expectations of Teachers
4. The Experimental Method	4. Observing Teachers' Expectations of Pupils
5. Descriptive Studies	5. Observing Pupils' Contacts with Pupils
6. Tools of Research	6. Observing Teachers' Contacts with Pupils
7. Descriptive Data Analysis	7. Observing the Classroom and its Materials
8. Inferential Data Analysis	8. Observing the School as an Organization
9. Writing the Report	Plus 29 Statistical Appendices
10. Historical Research	



A major strength of *Educational Research in Classrooms and Schools* is its widespread sampling of the enormous variety of techniques, instruments, and research findings in its particular area of concern — the classroom and school situation. Each of the sections presents the basic principles, gives examples of the major tests, methods, and materials involved, and gives a summary of research studies in the area. Each section also contains a useful bibliography and suggestions for further research. A particularly valuable feature for a beginner is the reproduction of excerpts (and sometimes all) of the actual instruments used. This does not take the place of reading the original studies, but it does give a clearer picture of how the data were obtained.

Perhaps the most positive aspect of *Research in Education* is that, unlike Cohen's book, it is a comprehensive treatment of the topics which invariably crop up in an introductory graduate course in educational research. With the exception of the chapter on Writing the Report (on which more later) all topics are covered fully and clearly, with a particularly effective treatment of the use of reference and research sources.

A questionable aspect of the Cohen volume is the apparent selectivity of studies and instruments included. For example, Cohen's discussion of machiavellianism centres on two instruments: Guterman's "Machiavellian Scale" and the sinister-sounding "Kiddie Mach Scale" of Nachamie. Neither of these instruments is mentioned in *Tests in Print*. This absence does not signify that these instruments are lacking in reliability or validity, but examination of the kind of data available through *Tests in Print* can go far to reassure a researcher that he is using the right tool for the job.

Some aspects of the treatment of readability in *Educational Research in Classrooms and Schools* bear closer examination. While opinions may vary as to their relative merits, there is little argument that the three most prominent readability measures are the Dale-Chall, the Flesch, and the Gunning Fog Index. None of these is included in the readability section. The four measures which *are* included consist of: the Spache formula (based on Dale-Chall); the McLaughlin SMOG Index (based on Gunning); the Mugford Readability Chart; and the Fry Readability Graph.

There is some inconsistency in Cohen's treatment of Spache's Readability Formula. Initially the formula is described as being based on the Dale-Chall (1948) 769-Word List, which is included in the text. Later it is suggested that those interested in further study of the Spache technique should consult the Dale-Chall (1948) 3000-Word List. Which is correct? Dale-Chall's 769-Word List actually comes from their 1931 paper on readability — not their 1948 paper.

Cohen recommends the Mugford Chart over the Fry Graph for speed and ease of application. Since the Mugford requires five separate operations and the use of 3 separate charts, compared to 3 steps and one graph for the Fry technique, his rationale is less than convincing, particularly when he states that there is as yet little evidence of validity for the Mugford technique.

Any major deficiencies in *Research in Education* would appear to involve omission rather than commission. Best includes report writing in his volume, but the presentation is far from comprehensive. Although the chapter is entitled "Writing the Report," in fact the chapter is exclusively concerned with writing a *thesis*, which is admittedly a form of research report, but a very specialized one. A major objection to Best's approach to the thesis format is that it presents only *one* alternative, featuring the heavily-footnoted citation technique which is mercifully being discarded by many institutions in favour of the APA format. A researcher tackling a research thesis for the first time (or even the second) has surely a

difficult enough passage without having to wrestle with a list of 40 footnote abbreviations such as *op cit*, *ibid*, *passim*, etc. Best is of course entitled to his preference, but his *readers* should at least be made aware of the alternatives.

As stated above, there is not much overlap in content between these two volumes. Two areas which *are* covered in both books are statistics and experimental method. At first glance Cohen's 29 statistical appendices would appear much more comprehensive in scope than Best's two chapters on descriptive and inferential statistics. On closer examination the coverage is seen to be approximately equivalent. Any author of an introductory research text is faced with the problem of how much statistical content can be usefully included. Even the 100 or so pages of text devoted to statistics by these two authors can scarcely expect to meet the requirements of the topic, when it is considered that a basic statistical treatment such as Ferguson's requires a 500-page volume.

In their treatment of experimental design, both Cohen and Best acknowledge their debt to Campbell and Stanley (who else?), but their adherence to the precepts of Campbell and Stanley differs considerably. Best uses the same notation and terminology as Campbell and Stanley and his treatment, while necessarily limited, is valid and straightforward.

The discussion of experimental design in *Educational Research in Classrooms and Schools* is less admirable. In his discussion of the Pre-Test, Post-Test, Control Group Design Cohen intimates that ideally the division into groups should be done at random, or the means and standard deviations of the groups may be compared. This is, of course, categorically wrong! Without the random assignment to groups, you have, not a true experiment, but a *quasi*-experiment, a term which does not even appear in Cohen's discussion.

Although library science treats the construction of a book index as a skilled technique calling for a high degree of expertise, this view appears to be foreign to book publishers. Many indices appear to have been compiled by the local office boy and these two volumes are no exception — both contain numerous errors in coverage and logic. Let us examine how the important research topic of reliability is handled by the index of each book. The Best index contains a single page entry, which turns out to be a brief definition of reliability. In fact the book also contains discussions of test-retest, equivalent forms, and split-half reliabilities, but these are included in the section on correlation methods and are *not* listed in the index. An entry for split-half reliability in Cohen's index refers to a single quoted coefficient for an instrument. Cohen's index also refers to the Spearman-Brown formula as a "reliability test."

It is difficult to compare these two works other than in terms of how well they appear to meet their stated objectives. In the case of *Research in Education* it may be said that this is the best attempt yet seen by this reviewer at an introductory text for a graduate course in educational research. The Cohen text certainly introduces a variety of topics, research methods, and data techniques for studying the classroom and school environments. A beginning researcher would be well advised, however, to keep in mind some of the questions raised above.

Andrew K. Clark  
*Department of Industrial and Vocational Education*  
*The University of Alberta*



THE CURRICULUM: THEORY AND PRACTICE. *By A. V. Kelly.* London: Harper and Row, 1977, 202 pp.

The spate of general books on curriculum continues unabated, a testimony to the enormous expansion of what the author of this volume calls “the curriculum industry” that has grown up on both sides of the Atlantic as planned change has become an educational norm. Kelly, who lectures in Curriculum Studies at Goldsmiths’ College, London, claims justification for this mercifully brief treatment (in comparison with most tomes in the field) in the fact that he has attempted an overview instead of dealing with a single aspect of Curriculum Theory. The capitalization is his and the implied assumption that there is a body of curriculum theory may be questionable. In addition, he has attempted to provide a dispassionate, if not an objective, explication of many theories and positions, ideological or otherwise, in the field to the end that the reader may gain “the understanding he needs to weigh them against each other.” The assumed aim is to assist the practising teacher to bridge the theory-practice gap. In this, and the related aims noted, Kelly has succeeded admirably.

Despite an eclectic approach, the volume does argue a specific case — that of the teacher’s role in curriculum development and the need to support him/her in that role. The argument is based on the familiar assumption, now almost the primary canon of practical curriculum development, that control of the process “will always rest in the hands of the teacher in the classroom.” Given this fact, the teacher must be supported and Kelly seeks to do that by making available some of the curriculum theories that abound in such a way as to promote understanding and thereby to contribute to more effective practice.

An opening chapter on curriculum planning focuses on issues relevant to the curriculum viewed as an “overall rationale for the educational programme” of any school together with general features of curriculum change and development. Throughout, the emphasis is on the problem, in Lawrence Stenhouse’s terms, of the relationship between intention and reality. Theory and practice must go hand in hand and curriculum studies must be based on an integrated, interdisciplinary approach.

Subsequent chapters deal respectively with curriculum objectives, content, integration, evaluation and the social context of curriculum development. Linear Tylerian development and objectives models are critically, concisely, and lucidly analyzed and alternative models, e.g., Eisner’s “expressive objectives,” are presented. Particularly cogent is a critique of Bloom’s taxonomy as lacking a concept of education.

Kelly is at his best in dealing with the issue of curriculum content. In thirty pages he manages to discuss a whole range of related issues: sources of cultural content as between “high” and “folk” cultures, knowledge transmitted for “its own sake,” Hirst’s view of the nature of knowledge and its relevance for a curriculum based on rationality, the view of knowledge as socially constructed and utilitarian views of knowledge and the curriculum, among many others. On the crucial question of who makes the selection of what is to be taught, Kelly, consistent with his overall thesis, concludes that inevitably this rests “with the professional judgement of the teacher.” Such judgement, itself inevitably subjective, must be informed and “firmly based on a full knowledge and understanding of the issues involved.” Understanding must rest on insight into the social context of curriculum development. The actual context discussed is naturally British, but relevant enough to the experience of Canadian teachers to be helpful, at the same time that it provides valuable information about various aspects of the overseas curriculum scene. Issues related to teacher-based curriculum development, to dissemination



and implementation, to evaluation, teacher accountability, and a common curriculum are all treated in straightforward fashion free of jargon and technical detail. Most of these themes are elaborated in separate chapters, the last, dealing with the issue of a “common curriculum,” being particularly relevant in a “back to basics” era.

For the nonspecialist at any level, this book is an excellent introduction to curriculum studies. It would be valuable as a core text in an undergraduate course or as a theoretical handbook in curriculum development projects and in-service work. A useful, up-to-date bibliography of nearly 150 articles, reports, and books is included.

George S. Tomkins  
*Centre for the Study of Curriculum and Instruction*  
*Faculty of Education*  
*The University of British Columbia*







## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.



# *ajer*

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXV    NUMBER 4    DECEMBER 1979

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    ●    EDMONTON



# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

Published quarterly in March, June, September, December by the  
Faculty of Education, The University of Alberta

## CONSULTING EDITORS

J. Britton  
*University of London*

J. Calam  
*The University of British Columbia*

M. Connelly  
*The Ontario Institute for  
Studies in Education*

K. De Clerck  
*State University of Ghent*

R. N. Evans  
*University of Illinois at  
Urbana-Champaign*

R. H. Farquhar  
*University of Saskatchewan*

E. Gagné  
*University of Ottawa*

G. Harman  
*University of Melbourne, Australia*

S. Hunka  
*The University of Alberta*

J. W. G. Ivany  
*Simon Fraser University*

D. A. MacIver  
*University of New Brunswick*

L. D. Nelson  
*The University of Alberta*

W. C. Nesbit  
*Memorial University of Newfoundland*

E. Pedersen  
*McGill University*

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

## FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith  
T. P. Atkinson  
N. C. Bhattacharya  
R. C. Bryce

W. T. Fagan  
H. W. Hodysh  
R. G. Martin (Chairman)

E. Miklos  
V. R. Nyberg  
C. H. Preitz

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

Volume XXV, Number 4

December, 1979

CONTENTS

Heuristic Strategies Utilized by High School Students .....	213
<i>L. Pereira-Mendoza</i>	
Newfoundland Dialect Interference in Fourth Grade Spelling .....	221
<i>L. Walker</i>	
The Development in Children and Adolescents of the Construct of Rational Numbers as Operators .....	234
<i>T. E. Kieren and B. Southwell</i>	
The Relationships of Language Concepts, Auditory Comprehension, Visual Perception, and Spatial Relations as Predictors to Reading Achievement in First Grade .....	248
<i>B. B. Adams and L. O. Ollila</i>	
Comparative Difficulty of Beginning Reading Vocabulary: Set II ....	259
<i>B. Horodezky</i>	
Teacher Socialization: The First Five Years .....	264
<i>J. McArthur</i>	
Student Teacher Telephone Conferencing with Satellite Maps as a Monitoring Device .....	275
<i>J. M. Kirman and J. Goldberg</i>	

PERSPECTIVES

Some Observations Regarding Value Added by Education .....	284
<i>R. N. Evans</i>	

ESSAY REVIEWS

Education in Atlantic Canada .....	288
<i>D. A. MacIver</i>	
Moral Education: The Risk of Oversimplification .....	294
<i>I. DeFaveri</i>	

(over)

BOOK REVIEWS

<i>Classroom Language: What Sort?</i> , J. Richards .....	307
Reviewed by Ellen M. Regan	
<i>An Experiment in Teacher Education</i> , A. Ross, D. McNamara and J. Whittaker .....	309
Reviewed by Lorne D. Stewart	
TABLE OF CONTENTS, VOLUME XXV, 1979 .....	310

FACULTY OF EDUCATION  
*The University of Alberta*



LIONEL PEREIRA-MENDOZA

*Memorial University of Newfoundland*

## Heuristic Strategies Utilized by High School Students

*This study investigated the heuristic strategies utilized by high school students (N = 20) to solve mathematical problems. The experimental session lasted up to one hour per subject, during which time the student attempted four problems, a maximum of fifteen minutes being allowed per problem. The subjects verbalized their ideas while solving the problems and the sessions were videotaped. Analysis of the protocols employed by the subjects indicated that examination of cases (including systematic cases and looking for a pattern in the results), symmetry, and analysis were the most common heuristics. Subjects tended to examine cases in many problems while employing the other heuristics in only one problem, indicating that generally the utilization of heuristics tended to be problem-dependent.*

In a world in which scientific and technological knowledge is growing exponentially, society is presented with the challenge of educating individuals who will be able to function in a continually changing environment. Thus, educators and psychologists are faced with the problem of identifying and teaching the competencies needed to deal with this changing world. At the center of these competencies are "strategies of inquiry." Heathers (1965) indicated that "problem-solving thinking or inquiry is generally considered to be the core of the educational process, and the chief mark of the educated person" (p. 2).

One approach taken by scholars has been to investigate the nature of the problem-solving process. Their efforts have resulted in analysis of various stages involved in problem-solving. For example, Bloom and Broder (1950) suggested the following four components: understanding the nature of the problem, understanding the ideas contained in the problem, general approach to the solution of problems, and attitude towards the solution of

---

Dr. Pereira-Mendoza is an associate professor of mathematics education at Memorial University, St. John's, Newfoundland, Canada. He received his undergraduate education in England and his doctoral work was undertaken at the University of British Columbia. His primary research interest is in the area of mathematical problem-solving, with particular reference to the heuristics utilized in solving novel problems.

problems. Garry and Kingsley (1970, p. 464) suggested three steps in problem-solving: the search phase, the functional solution phase, and verification of the final solution. Gagné (1966, p. 138) included the stages of defining the problem by distinguishing essential features, searching for and formulating hypotheses, and verifying the solution in his analysis of the problem-solving process. While such analyses provide an overall framework for examining the nature of the problem-solving process, they do not include specific techniques that can be applied by the learner to make discoveries about novel mathematical problems. These specific techniques can be defined as heuristics.

*Mathematical heuristics* are procedures which are used for the purpose of discovering mathematical relationships in a novel problem, but whose application does not guarantee success. For the mathematics educator, the work of Polya (1954a, 1954b, 1957, 1962, 1965) has provided the major source of heuristic strategies. Polya's heuristic framework is stated in the form of specific questions that can be posed in order to try to solve problems. Some of these questions are: "Do you know a related problem? . . . Could you imagine a more accessible related problem? A more general problem? A more special problem? An analogous problem? Could you solve part of the problem? Draw a figure . . ." (Polya, 1957, pp. xvi-xvii). MacPherson (see Blake, 1976) refined the Polya framework to produce a set of heuristics, and it is this set that formed the framework for this research. Some of the heuristics in this set are presented under a subsequent heading, *Coding Form*.

### *Research and Purpose*

Much of the current interest in clinical research on heuristic strategies in the area of mathematics education was initially generated by Kilpatrick's study (1967) which utilized a coding scheme based on Polya's heuristic framework. While other studies (e.g., Lucas, 1972; Webb, 1975; Kantowski, 1977) tended to employ distinct but not disjoint sets of heuristic strategies and coding schemes, it seems reasonable to conclude that most researchers utilized the Polya framework and Kilpatrick's coding scheme as the initial basis for their research, each researcher modifying them in view of additional research findings. The overall implications of such research seem to be clear, namely that students do utilize and can be taught to utilize various heuristic strategies. For example, in many of these studies, it was found that students drew figures, made tables, looked for patterns, and tried trial and error (random cases) when solving problems. In these studies, algebraic or word problems were often employed to generate heuristic strategies.

The purpose of this study is to utilize the heuristics outlined by MacPherson as a basis for investigating the strategies employed by students in solving geometric problems. Furthermore, this study allows the students to utilize physical materials when solving problems, a departure from the other studies. The specific questions investigated were:

1. Without specific training in heuristics, what heuristic strategies are employed by high school students when solving novel mathematical problems?
2. What are the most commonly utilized heuristics?

### Procedure

A sample consisting of 20 high school students (ten in grade 9 and ten in grade 11) was selected from the local schools. The subjects were all in the honours mathematics program and had been identified by their teachers as the better problem solvers.

Each subject came to the university at a different time and was taken to a room where the videotaping equipment had been set up. Subjects were given four problems to solve, with a random number table being utilized to determine the order of the problems for each individual. Prior to taping, they were given a general orientation towards the equipment and an opportunity to ask any general questions regarding the taping. It was explained that they were expected to think aloud while solving the problems. Finally, the subjects were informed of the maximum of 15 minutes per problem, and that at the end of this period or on completion of the problem (whichever came first), they would continue on to the next problem. It should be noted that a problem was considered complete when the subject obtained a solution, whether or not that solution was correct. No comments or advice were given during the videotaping, except for the occasional reminder to think aloud.

### Problems

The four problems are listed below. In addition to the materials appropriate to each problem, subjects were provided with writing paper, pens, and a ruler.

*The Chinese Tangram Puzzle.* The seven pieces were placed in an envelope in front of the subject, together with an index card on which was written:

Arrange all the pieces in the envelope to form a square.

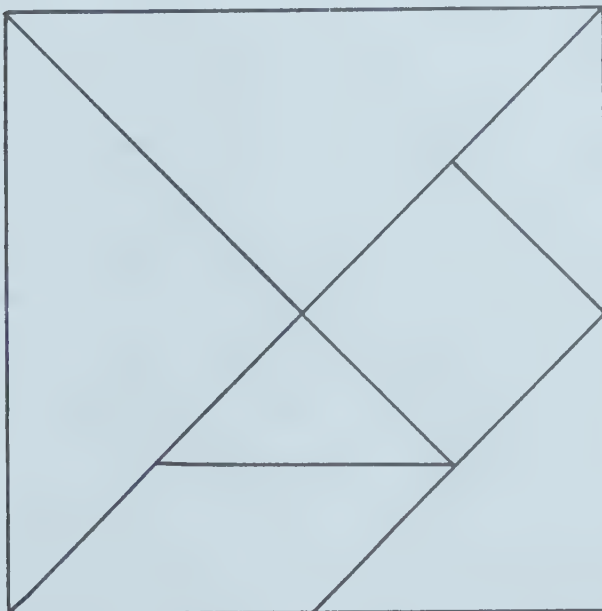


Figure 1. Tangram solution.

*1,000 Squares Problem.* The subject was presented with a plastic bucket containing a large number of congruent squares, together with an index card on which was written:



Given 1,000 congruent squares how can they be arranged in a plane so that the total perimeter of the figure formed is a minimum? Find the perimeter.

*Ten Lines Problem.* The subject was presented with a geoboard, elastic bands, and an index card on which was written:

Find the maximum number of points of intersection of ten lines in a plane.

*5 x 5 Rectangle Problem.* This problem was an adaptation based on the classic problem: How many squares on a chessboard? The subject was presented with a ruled piece of card (approximately 25 cm by 25 cm) divided into 25 congruent squares, together with an index card on which was written:

Find the total number of rectangles on this board.

### *Coding Form*

The coding form employed was basically that developed by Blake (1976); a detailed discussion can be found in his dissertation. Since this study was concerned with the heuristics utilized by students, only that aspect of the form is discussed. The coding form enabled the coder to enter both heuristics employed and the order in which they were applied. The form contained twelve heuristics including: *smoothing*—when the problem is altered in order to obtain some isomorphism between the problem and a mathematical system; *analysis*—breaking the problem into subproblems; *cases*—renaming the variables in a problem as constants (includes all cases, critical and systematic cases); *exclusion*—a variable is excluded from the problem; and *symmetry*—making use of intrinsic or constructed symmetry in a problem. A detailed account of the twelve heuristics, together with examples, can be found in Blake (1976, pp. 7-18).

### *Analysis of the Data*

Using videotapes retained from the pilot study, the judges were trained in the coding procedure. At the end of the training sessions, the analysis of the data tapes was undertaken. The judges coded each tape separately, then met to discuss the results. Only when all judges agreed that there was evidence of a given strategy was that evidence utilized in the analysis.

The most commonly utilized heuristic was one of the “cases family.” It is important to note that while some subjects examined cases in a systematic manner, not all of them used the data to try and formulate a pattern. For example, in the Ten Lines problem, some subjects tried 2, 3, 4, and 5 lines in a systematic manner (obtaining the correct answers), but then concluded that the problem was getting complex and started again. They made no attempt to search for a pattern in their results. Some subjects started to search for a pattern after only three cases while others required five or six cases.

The two examples of critical cases both occurred in the 1,000 Squares problem. In each case, the subject made comments to the effect that if all the squares were separated, the perimeter would be 4,000 and hence the answer must be less than that. Thus, the subjects utilized this as a basis for establishing a bound for the answer.

Attempts to solve the problems employing symmetry occurred most

TABLE 1  
USE OF HEURISTICS<sup>a</sup>

Heuristic	No. of subjects who utilized the heuristic at least once when solving the four problems (Maximum 20)
Cases	14 (utilized at least one of the various facets of this heuristic---see below)
(a) All	4
(b) Critical	2
(c) Systematic	10
(d) Systematic Plus Pattern	6
Analysis	6
Symmetry	8
Analogy	1
Inverse Deduction	1

<sup>a</sup> Note that the use of random cases is not included in this list. The nature of the problems, such as the tangram problem, meant that the subject either had to do nothing or manipulate the equipment in some manner. As such, it was felt that random manipulation could not be considered an heuristic, but rather a function of these specific problems. In other problem situations it could be classified as an heuristic.

frequently in the Tangram and Ten Lines problem (see Figure 2). For example, in the Ten Lines problem, the subjects would *continually* try to form the square by organizing the pieces in a symmetrical manner. It is significant to note that most of the subjects who utilized symmetry continued with a fixation effect on this heuristic. For example, once having decided that the solution to the Ten Lines problem was symmetrical, they would continue with a symmetrical approach, usually resulting in the conclusion that the maximum number of points of intersection was 25. Symmetry was also noted in the 5 x 5 Rectangle problem where a subject would count the 2 x 3 rectangles, say, horizontally and conclude that there were the same number vertically.

The only other heuristic that occurred with any regularity was that of analysis. This heuristic was employed most frequently in the Tangram and 5 x 5 Rectangle problems. In the Tangram problem, the subjects said that they could solve the problem by making two rectangles, or two congruent right triangles, or in one case, four congruent right triangles. They then proceeded to try to manufacture the appropriate figures. In the 5 x 5 Rectangle problem, the subjects tried to solve the problem by breaking it into two subproblems, namely the number of squares and the number of rectangles that were not square. In order for analysis to be coded it was necessary that a subject had *both* made a comment indicating a breaking of the problems into subproblems and followed through by applying the approach suggested by the analysis to try to solve the problem.

It is important to note that a subject who applied an heuristic to one particular problem did not always utilize that heuristic in another problem.

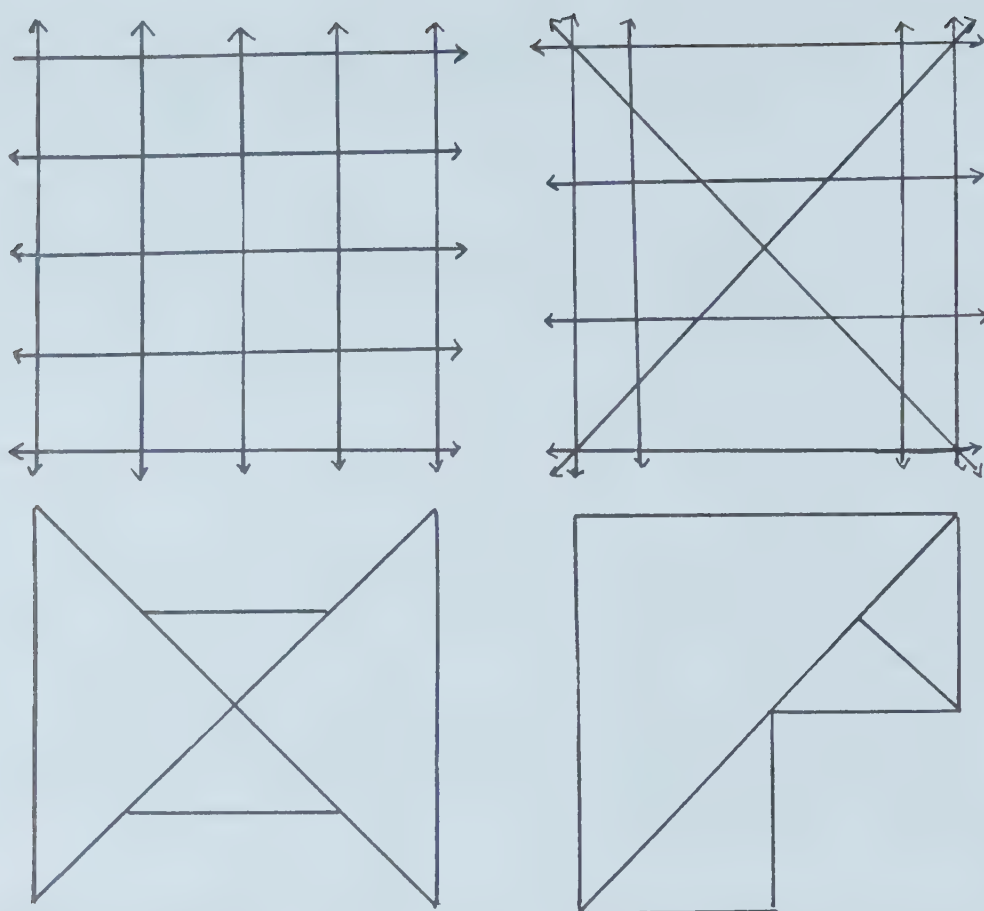


Figure 2. Symmetric arrangement for the Tangram and Ten Lines Problem.

For example, a subject who tried symmetry in the Ten Lines problem did not necessarily try the heuristic when solving the Tangram problem. It appeared that the utilization of an heuristic was *very* dependent on a subject's perception of the problem. This result is consistent with Webb's (1975) conclusion that heuristics were problem-dependent.

Of the 20 subjects involved in the study, 16 applied at least one heuristic during the experimental session. The problem-solving approach utilized by the other four can best be described as totally random "trial and error." Most subjects who applied systematic cases (either with or without searching for a pattern) did so in all the problems, except for the Tangram problem. It appeared that the subjects were willing to try this heuristic in a variety of problem settings. The same conclusion cannot be drawn regarding analysis and symmetry. For example, although symmetry was utilized by a total of 8 subjects, only 2 of them applied the heuristic to two problems.

Table 2 indicates that only three subjects applied three heuristics during their experimental sessions; ten applied two heuristics, usually one of the cases family and one other heuristic. Of the three subjects who utilized only one heuristic, two applied cases and the other applied symmetry,

Finally, one particular aspect of the problem-solving process not discussed so far is the role of physical materials. Analysis of the videotapes indicated that the grade 9 subjects utilized the physical materials far more than the grade 11 subjects. For example, in the Ten Lines problem, *all* the



TABLE 2  
NUMBER OF DIFFERENT HEURISTICS UTILIZED BY STUDENTS<sup>a</sup>

	Number of heuristics			
	3	2	1	0
Grade 9	1	4	2	3
Grade 11	2	6	1	1

<sup>a</sup> Subjects who utilized one or more of the components of the cases family have been given credit for one heuristic.

grade 9 subjects used the geoboard, while five of the grade 11 subjects immediately pushed the board aside and started drawing figures (two of the five who pushed the board aside decided to use it later in the problem). Only the Tangram problem produced no differences, with all subjects manipulating the pieces.

Summary of Results

Evidence of five heuristics was found, namely cases (all, critical, and systematic—both with and without searching for a pattern), analysis, analogy, symmetry, and inverse deduction. However, only one example of each of analogy and inverse deduction was found. The most commonly utilized heuristic was cases, particularly systematic cases, followed by symmetry and analysis. No subjects utilized more than three heuristics during the experimental session, with most subjects (10) employing two heuristics. Finally, the grade 9 subjects relied more heavily on the physical materials than the grade 11 subjects, although both groups utilized the materials.

There are many possible directions for further research resulting from this study. First, as was indicated earlier, the utilization of heuristics appears to be problem-dependent, and the study should be extended to other problems. Second, the sample selected for this study (and for many heuristic process studies) consisted of “better” problem solvers, and the investigator feels that the future studies should include subjects who are both “average” and “below average” mathematical problem solvers. Finally, an extension to include subjects in both elementary-junior high school and college could provide evidence concerning heuristic development during the complete schooling period. Such a study could provide evidence as to whether heuristics appear at different stages of schooling.

References

Blake, R. The effect of problem context on the problem solving ability of field independent-dependent students: A clinical study. Unpublished doctoral dissertation, The University of British Columbia, 1976.

Bloom, B. S., & Broder, L. J. Problem-solving processes of college students. *Supplementary Educational Monographs*, Chicago: University of Chicago Press, 1950, No. 73.

Gagné, R. M. Human problem solving: Internal and external events. In B. Kleinmuntz (Ed.), *Problem solving: Research, method, and theory*. New York: Wiley, 1966.

- Garry, R., & Kingsley, H. L. *The nature and conditions of learning* (3rd ed.). New Jersey: Prentice-Hall, 1970.
- Heathers, G. *Four process goals of education*. Center for Research and Development on Educational Differences, Harvard Graduate School, 1965.
- Kantowski, M. G. Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education*, 1977, 5, 163-180.
- Kilpatrick, J. Analyzing the solution of word problems in mathematics: An exploratory study. (Doctoral dissertation, Stanford University.) Ann Arbor, Michigan: University Microfilms, 1967. No. 68-06442.
- Lucas, J. F. An exploratory study of the diagnostic teaching of heuristic problem-solving strategies in calculus. (Doctoral dissertation, University of Wisconsin.) Ann Arbor, Michigan: University Microfilms, 1972. No. 72-15368.
- Polya, G. *Mathematics and plausible reasoning, Volume I. Induction and analogy in mathematics*. New Jersey: Princeton University Press, 1954. (a)
- Polya, G. *Mathematics and plausible reasoning, Volume II. Patterns of plausible inference*. New Jersey: Princeton University Press, 1954. (b)
- Polya, G. *How to solve it*. New York: Doubleday Anchor Books, 1957.
- Polya, G. *Mathematical discovery, Volume I*. New York: Wiley, 1962.
- Polya, G. *Mathematical discovery, Volume II*. New York: Wiley, 1965.
- Webb, N. An exploration of mathematical problem-solving processes. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., March, 1975.

LAURENCE WALKER

Mount Saint Vincent University

## Newfoundland Dialect Interference in Fourth Grade Spelling\*

*In this study, the influence of selected phonological features of the dialect spoken along the northeast coast of Newfoundland on fourth grade spelling performance was investigated. A test based on words whose pronunciation was affected by these features was administered to approximately 400 subjects from the dialect-speaking area and to a similar number of subjects in Nova Scotia. Using a standardized spelling test score as a covariate, analyses of the scores of 100-subject samples showed a significant effect in favour of the control group. Analysis of composition samples revealed patterns of misspelling that confirmed the major finding that phonological features of dialect speech did influence spelling errors.*

Chomsky and Hallé (1968) claimed that English orthography is a near-optimal system for representing the spoken language. Part of the claim is based on the assumed power of a single, conventional system of spelling to represent, uniquely and without prejudice, all the various dialects of English. For the skilled reader it is of tremendous convenience to have only one writing system throughout the English-speaking world. Besides there is no convincing evidence, in spite of a large body of research, that for the beginning reader who speaks a nonstandard dialect, a written code which does not conform closely to his own dialect speech is a significant handicap. Less attention, however, has been paid to the question of whether nonstandard dialect speech interferes with the acquisition of the ability to spell conventionally. In the case of the beginning reader, recognition of a word involves going from a defined graphic stimulus to one of a range of permitted oral responses; for example, the written word "tomato" can be

---

Dr. Walker taught in public schools in Britain, Alberta, and Newfoundland before completing a doctorate at the University of Alberta in reading education. After three years as an assistant professor at Memorial University, Dr. Walker is now associate professor and Chairman of the Education Department at Mount Saint Vincent University, Halifax, Nova Scotia. The effect of dialect on spelling is part of his general research interest in children's use of written language.

\* This study was funded by a grant from the Canada Council.



pronounced either with the American/*tomeyto/* or the British/*tomahto/*. In other words dialect responses to printed words are psychologically valid if not always pedagogically so. For the naive speller, on the other hand, the task is one of going from one of a range of pronunciations to a single printed form. On the face of it, therefore, the nonstandard dialect speaker might be assumed to have more difficulty with spelling than with reading.

However, as teachers and spelling authorities have long recognized, the relationships between English sounds and spelling are not simple even for the speaker of the most conventional dialect. All spellers are faced with the difficulty of deciding which of several alternative representations of a particular sound or sequence of sounds is the conventional spelling. For the nonstandard dialect speaker the question is whether there is an additional difficulty in his making this choice owing to further divergencies between his dialect pronunciation and that of more standard dialects. The term "nonstandard dialect" is used to refer to a regional or social class variety of spoken English which has a number of distinguishing phonological, lexical, and syntactic features that set it off from other dialects. It is recognized that there is no single form of standard spoken English, and that so-called standard English, the variety of English spoken by the educated and influential members of a community, varies, mainly phonologically, from one geographical region to another.

### *The Role of Speech Sounds in Spelling*

The larger question which underlies the issue of dialect interference in spelling is the extent to which the speller utilizes the sound information in the spoken word to be spelled. Traditionally, English was considered to be too irregular in its sound-letter correspondences for pronunciation to be a very useful aid to spelling. From an orthographic point of view, this perception was challenged by computer studies of large sets of words (Hanna, Hanna, Hodges, & Rudorf, 1966; Venezky, 1967). These studies revealed that when variables such as morpheme boundary, form class, stress, and graphemic environment are taken into account, English orthography demonstrates considerable regularity of sound-spelling correspondence. Venezky's (1967) conclusion that a morphophonemic level intervenes between the speech sound and its graphic representation was supported by a similar theory proposed by Chomsky and Hallé and applied to examples of derived word spellings by C. Chomsky (1970). This theory proposed that the phonological system includes an abstract lexical level from which surface phonetic pronunciations are generated, and that the spelling system represents this abstract level in a relatively consistent manner rather than the phonetic level in an irregular manner. For example, in "extreme" and "extremity" the second "e" represents two different phonemes; yet the spelling preserves the lexical relationship between the two words.

In a recent review of the issue of sound-spelling correspondences, Cahen, Craun and Johnson (1971) concluded that "overall the evidence indicates that phoneme-grapheme correspondences of some consistency do exist in English" (p. 297).

*Spelling Processes and Dialect Interference*

Having established that the English spelling system is related to sound, the question then becomes one of whether the speller makes use of the potential information in this relationship in order to help spell words. The extent to which he does and the manner in which he does is directly relevant to the question of dialect interference in spelling.

Theoretical analyses of the spelling process have been proposed by Simon and Simon (1973) and Peters (1970). Simon and Simon (1973) acknowledged that most people probably spell most words that are highly familiar to them by a direct recall method. Otherwise, they proposed, phonemic information is used in two ways: through stored phoneme-grapheme associations supplemented by rules governing the selection of different alternatives, and through a generate-test method analogous to the sounding out process in word recognition in reading (p. 119). This last strategy involves generating a list of possible spellings for an unfamiliar word and then using a visual recognition process, derived from experience with reading words, to select the appropriate form. The authors reported the use of a computer algorithm programmed to use both possible graphemic representations of phonemes and word recognition information to spell a body of words. The misspellings resulting from this computer program closely resembled those made by a group of grade 4 students on words that they had not previously studied. Simon and Simon (1973) concluded that this finding provided empirical support for the generate-test hypothesis as an explanation of "much of the spelling behaviour of fourth graders" (p. 135).

Peters (1970) proposed a similar procedure in which one of several possible orthographic representations is selected by the speller. However, unlike Simon and Simon, she based her selection procedure on serial probability information, or the awareness on the part of the speller of the most likely letter sequences permitted in the orthography. For example, "lodj" as a phonemic spelling contains the sequence "dj" which has a low or even zero probability in English orthography. Therefore spellers would likely reject it in favour of "lodge" whose "dge" sequence has a relatively high probability of occurrence. On the basis of the results of a letter sequence generalization task, included in a larger study of predictors of spelling achievement in 9-10 year olds, Peters (1970) claimed support for her hypothesis. However, no validity or reliability information on the task was offered in the report. Wallach (1963), earlier testing a similar hypothesis, concluded that the results of his study of fifth-grade spelling of different patterns of nonsense words were consistent with the proposition that good spellers as opposed to poor spellers are "learning a general property of English words, namely the transitional probability governing the sequential arrangement of letters in . . . words" (p. 61).

Thus, in approaching the question of dialect interference in spelling, the implications of three theoretical positions can be considered: the theory, more orthographic than psychological, that spelling represents an abstract lexical level in the phonological system; the generate-test theory; and the serial probability theory. Since the abstract lexical level theory is not addressed by this study, its implication for direct interference in spelling



will not be discussed here. Such a discussion is available in Brengelman (1970).

The generate-test theory of the spelling process would seem to have strong implications for the effects of dialect on spelling. In the stage where possible spellings are being generated, the set of possibilities produced by a dialect speaker for certain words might be expected to be different from the set produced by a standard-English speaker. For example, a speaker for whom a reduction rule deletes /d/ in word final consonant clusters might generate a list for / hænd / that would include "han." However, at the test stage "han" would likely be rejected since it would not be a word ever encountered in a reading situation; and, provided "hand" was among the options generated, the correct spelling would likely be selected. This theory would suggest that dialect speakers are likely to display spelling difficulties when their pronunciation of a word produces another word in standard English. For example, a speaker who pronounces "pig" like "peg" or "born" like "barn" may generate among his lists of alternative spellings forms that, while recognizable as familiar words, do not conventionally represent the appropriate lexical targets. Based on the generate-test theory, a dialect-based spelling error would require two conditions: first that the word to be spelled was not a familiar spelling word (in the case where it was, the word would presumably be spelled directly) and, second, that the word written as a misspelling was familiar in a reading or word recognition sense.

The serial probability theory would predict a wider range of dialect misspellings since the test stage in this view of the spelling processes relies on the likelihood of a sequence of letters being a permissible pattern in English. In this case "han" as a spelling for / hænd / would be predictable since the letter sequence is a common one.

Research into the influence of dialect on spelling performance has generally been concerned with testing the hypothesis that dialect does interfere rather than testing theoretical explanations of how any interference might occur. Several studies have provided evidence that some interference exists.

Graham and Rudolf (1970) studied grade 6 students in three geographical areas of the United States and found that dialect speech did appear to influence spelling. Wolfram and Whiteman (1971) found examples of misspelled words in compositions written by tenth-grade Black students that appeared to be related to dialect speech. Kligman, Cronnell, and Verna (1972) compared Black English-speaking grade 2 pupils with a sample of standard English-speaking pupils. When spelling errors on a multiple-choice test were divided into dialect-based and nondialect-based misspellings, they found a significant interaction between speaker group and error type. Nineteen per cent of the Black "dialect" speakers' errors were dialect-based as opposed to 12 per cent of the standard English speakers'. The study also revealed that certain dialect features were more prone to be implicated in misspellings than others. In a follow-up study, Groff (1978) found that by grades 4, 5, and 6, significantly fewer errors on the same test were dialect-related and that only a few of the dialect features seemed to be causing interference with spelling by the time pupils reached the middle grades. The test items which continued to produce dialect-related misspellings were:



“still,” “walks,” “streets,” and “nests.” It is interesting to note that the misspellings of these words which would be considered to be dialect-based would be respectively: “steel” or “steal,” “walk,” “street,” and “nest.” The fact that real word misspellings were the most persistent errors would be congruent with the generate-test theory proposed by Simon and Simon (1973). However, other test items such as “your” (you) and “its” (it) did not apparently continue to produce errors based on pronunciation.

In the case of other nonstandard dialects, a study by Boiarsky (1969) found that features of an Appalachian dialect were related to misspellings significantly more frequently in the spelling of tenth-grade students from rural Appalachia than in the spelling of students in a Philadelphia sample. An informal study in Newfoundland by Walker and Paddock (1975) found that of 253 spelling errors identified in stories written by a sample of 57 pupils from grades 3 to 6 in two rural schools, 62 errors (almost 25 per cent) appeared to be related to features of the local dialect.

In summary it seems fair to say that consistent but complex phoneme-grapheme relationships exist in English orthography. Different theories of the spelling process have suggested ways in which the speller learns to use the information provided by these relationships and the hypothesis could be derived from these theories that the speaker of a nonstandard dialect would have additional difficulties in learning to spell conventionally. Moreover, there is some evidence from studies of Black English speakers and of Appalachian dialect speakers that spelling errors are related to certain characteristics of divergent speech patterns. The present study sought to examine the question of interference in spelling as it applies to another dialect of English, that spoken in the Canadian province of Newfoundland. (The conditions under which the dialect developed are briefly described elsewhere [Walker, 1975].)

### *The Study*

The purpose of the present study was to test the hypothesis that certain phonological features of a Newfoundland dialect cause spelling interference at the grade 4 level.

#### *Dialect Spelling Test*

The major source of data was a test of dialect spelling. Ten phonological features of the dialect spoken on the northeast coast of Newfoundland were selected (Paddock, 1977; personal communication, 1975). (Table 4 lists these features using Trager and Smith, 1951, notations.)

Word lists from three spelling series used in Newfoundland and the Province of Nova Scotia were then examined. All words whose pronunciation could be affected by one of the features shown in Table 4 and that were included in the grade 2-4 levels in these three series were listed. A random sample of the words was taken to produce a set of 36 test items with each of the 10 categories of features being represented in roughly the proportion revealed by the total population of words on the lists.

The 36 words were then put into random order and a context sentence written for each one. The words and sentences were tape recorded as an oral spelling test. A female, standard Canadian English-speaker did the

recording on good-quality tape in a university audiovisual studio. Each test word was pronounced, followed by the context sentence and the pronunciation of the word a second time. Sufficient time was allowed between the items for students to write the word down. As the result of a pilot test in one rural grade 4 classroom, two nondiscriminatory test items were replaced by other words from the same categories in the original lists. (The 36 test items are shown in Table 4.)

*Population and Sample*

The dialect-spelling test was administered to a sample of Newfoundland and Nova Scotia grade 4 students. The target population consisted of grade 4 students in three school districts on the northeast coast of Newfoundland. A control sample was drawn from the grade 4 students in two jurisdictions in Nova Scotia. Both areas contained rural populations engaged primarily in the fishing and lumbering industries. In both areas all schools containing grade 4 students, adjacent to salt water, and connected to the mainland by a road were identified. Table 1 shows details of the populations and samples in the two provinces. The Newfoundland schools were reduced by stratified random selection to a sample of 13 which reflected, in terms of size of school, the larger qualifying population of schools. Sixteen of the 17 Nova Scotia schools were retained in the control sample; the seventeenth, a small isolated school, could not be fitted into the data collection schedule. Table 1 also shows the percentage of grade 4 students in three types of schools: large schools, defined as those having more than 40 fourth graders; medium-sized schools with between 20 and 40; and small schools with fewer than 20. Also shown is the percentage of fourth grade students in multigrade classrooms. The figures show that the Nova Scotia sample contained higher percentages of students in medium-sized and small schools and in multigrade classrooms than the Newfoundland sample. Another source of difference was the fact that, although the economies of the two regions were based on the same

TABLE 1  
POPULATION AND SAMPLE

	Newfoundland	Nova Scotia
<u>Population</u>		
Schools	28	17
Grade 4 students	900	450
<u>Sample</u>		
Schools	13	16
Grade 4 students tested	391	430
<u>Final Sample</u>	100	100
<u>Percentage of students in:</u>		
large schools	62	38
medium schools	26	45
small schools	12	17
multigrade classes	26	34



industries, the Nova Scotia control sample was drawn from a more prosperous population than the Newfoundland sample.

Although complete data were collected from 391 Newfoundland students and 430 Nova Scotia students, analysis was carried out on randomly-selected 100 students from each sample.

### *Data Collection*

Apart from the dialect-spelling test, two other sets of data were collected. To obtain a measure of general spelling performance, the Level 2, Form A Spelling Sub-test of the *California Achievement Tests* (Tiegs & Clark, 1970) was administered. This test requires students to identify a misspelled word from sets of 4 words. None of the misspelled words in the test appeared to have any coincidental connection with the phonological features used to derive the dialect-spelling test.

As a source of data on dialect intrusion into spelling under ordinary writing conditions, a sample of free composition was collected from each subject. A writing task was designed and followed carefully in each classroom. Some soap bubbles were blown and the students allowed to watch, touch, smell, and taste them. Then a discussion was encouraged followed by a short story about a bubble, read aloud to the class. At this point the students were asked to make up and write their own stories about a bubble. No help was given with spelling; students who asked for help were told to do their best. No time limit was imposed on the writing.

All data were collected within a ten-day period in April, 1976. Collection was carried out by two research assistants, one in Newfoundland and the other in Nova Scotia. Both were females with teaching qualifications and experience. Data were collected in two sessions with each class.

### *Data Analysis*

Analysis of covariance was used to test for significant differences between the two samples on the dialect-spelling test. Scores on the standardized spelling test were used as the covariate.

All misspellings on the free compositions were tallied and those which appeared to be related to any of the 10 phonological dialect features involved in the study were noted. The proportions of misspelled words that could be attributed to these dialect features were compared using a *t* test.

## *Results*

### *Dialect-Spelling Test Scores*

The results of the analysis of covariance of the scores on this 36-item test are shown in Table 2. The means and standard deviations for dialect-spelling scores are shown in Table 3.

Table 2 shows that there was a significant difference between the two samples on spelling ability as measured by a standardized test. The mean score for the Newfoundland sample was 16.46 with a standard deviation of 5.37; for the Nova Scotia sample the mean was 18.21 with a standard deviation of 4.14. Clearly, therefore, the samples were unequal with respect to general spelling performance, making it necessary to control for spelling



TABLE 2  
ANALYSIS OF COVARIANCE: DIALECT-SPELLING TEST SCORES BY  
PROVINCE WITH STANDARDIZED SPELLING TEST SCORES  
AS A COVARIATE

Source	SS	df	MS	F	p
Covariate	7067.26	1	7067.26	252.86	.001
Province	875.74	1	875.74	31.33	.001
Residual	5506.02	197	27.95		

TABLE 3  
DIALECT SPELLING TEST SCORES: MEANS AND STANDARD DEVIATIONS

Province	N	Mean	SD
Newfoundland	100	24.44	9.24
Nova Scotia	100	30.70	5.54

ability. The Pearson *r* correlation for the total combined sample (*n* = 200) between the standardized spelling scores and the scores on the dialect spelling test was .72.

The main finding shown in Table 2 is that, over and above general spelling ability, there was a significant difference between the two groups on the dialect-spelling test. Table 3 shows that this difference favoured the control sample. In other words, on a spelling test made up of words deliberately chosen because logically they would be susceptible to misspelling by Newfoundland dialect speakers, the Newfoundland sample performed significantly lower than the Nova Scotia sample even when differences in general spelling ability were controlled. The design of the study suggests that this finding can be explained by the dialect differences between the two samples and that the phonological features of the nonstandard Newfoundland dialect did contribute to the spelling difficulties of the Newfoundland fourth graders.

Table 4 shows details of the performance of each group on the 36 dialect-spelling test items and overall on the 10 feature categories. Eleven of the test items produced a substantial number of misspellings that were predictable from the dialect features involved: "poor," "pool," "bill," "still," "kid," "fill," "thick," "both," "stem," "picked," and "spelled." Six of the ten phonological categories were responsible for substantial percentages of dialect-based misspellings. However, in the case of the one-word category /uw/ → { /ɔ/ /<sub>Δ</sub>/ } /—r the control group produced a higher percentage of errors that were apparently dialect-based than the target group, albeit with a smaller number of errors being involved. In the other five sensitive categories the control group's

TABLE 4  
TOTAL SPELLING ERRORS AND DIALECT SPELLING ERRORS  
BY PHONOLOGICAL CATEGORY

Category/word	Newfoundland (n = 100)		Nova Scotia (n = 100)	
	Total Misspellings	Dialect Misspellings	Total Misspellings	Dialect Misspellings
1. /ɔ/ + /æ/ / — r				
corn (carn)	17	2	5	0
born (barn)	26	0	7	0
morning (marning)	35	0	16	0
corner (carner)	33	0	25	0
Category percentage		1.8%		0%
2. /uw/ + { /ɔ/ / /ʌ/ } / — r				
poor (pore, pur)	24	13	11	8
Category percentage		54.2%		72.7%
3. /iy/ + /e/ / — r				
deer (dare)	23	1	13	1
Category percentage		4.3%		7.7%
4. /uw/ + { /u/ / /o/ } / — l				
pool (pull, poll)	24	17	6	2
cool (cull, coll)	48	5	12	3
Category percentage		30.6%		27.8%
5. /ɪ/ + { /iy/ / /e/ } / — l				
bill (beel, bell)	42	26	3	2
still (steel, stell)	38	21	8	1
kill (keel, kell)	42	12	0	0
fill (feel, fell)	36	24	11	1
Category percentage		52.5%		18.2%
6. /θ/ + /t/				
thick (tick)	29	12	9	1
everything (everyting)	57	3	31	1
both (boat)	39	8	22	5
teeth (teat)	23	4	11	1
Category percentage		18.2%		10.9%
7. /θ/ + /d/				
these (dese)	45	0	19	0
grandmother (grandmodder)	30	1	13	0
Category percentage		1.3%		0%
8. /h/ ↔ zero				
evening (hevening)	24	0	20	0
afternoon (hafternoon)	25	0	7	0
ice (hice)	2	0	1	0
hay (ay)	28	1	4	0
hang (ang)	27	1	12	0
hole (ole)	51	2	13	0
Category percentage		2.5%		0%
9. /e/ + /ɪ/				
led (lid)	48	0	60	1
pet (pit)	4	1	1	0
letter (litter)	27	4	11	2
yesterday (yisterday)	34	0	14	0
stem (stim)	46	14	23	4
shell (shill)	21	4	10	0
Category percentage		12.8%		5.9%
10. consonant cluster reduction				
picked (pick)	47	32	25	14
spelled (spell)	42	15	14	7
sold (sole)	13	1	16	1
pound (poun)	41	5	21	0
salt (sal)	31	0	22	1
coast (cose)	34	2	31	0
Category percentage		17.9%		17.8%

percentage of dialect-based spelling errors was lower than the Newfoundland sample's and involved considerably fewer misspellings. Within the consonant cluster reduction category, the greatest effect seemed to involve the -ed form of the verb. There was little or no evidence that the free alternations between / ɔr / and / ær /, / iyr / and / er /, and / ɔ̃ / and / d /, and the difference in the Newfoundland use of initial / h / affected spelling performance.

TABLE 5  
SPELLING IN FREE COMPOSITION

	Newfoundland (n = 100)		Nova Scotia (n = 100)	
	Mean	S.D.	Mean	S.D.
Composition length (words)	78.3	44.6	114.2	50.8
Misspelled words	8.3	4.6	7.3	5.8
Dialect-related misspelled words	1.12	1.64	.39	.72
Percentage of misspelled words that were dialect-related	12		5	

*Free Writing Misspellings*

Table 5 shows the results of the descriptive analysis of the free compositions written by the two samples. The Newfoundland compositions were on the average shorter in length; they contained more misspelled words; and a larger percentage of their misspelled words could be related to the phonological features of the dialect. A *t* test of the difference between the two samples on the percentage of spelling errors attributable to dialect was significant beyond the .001 level (*t* = 3.76, *df* = 198). These findings indicate that the Newfoundland students made spelling errors in their free compositions that were significantly more likely to be relatable to phonological features of their dialect speech than the spelling errors of the control sample.

*Discussion*

Clearly the Newfoundland sample experienced more spelling difficulties with the dialect-spelling test than did the Nova Scotia sample, even when spelling ability, as measured by a standardized spelling test, was statistically controlled. Since the items on the dialect-spelling test were selected on the basis of their relationships to certain phonological features of the dialect speech of the Newfoundland sample, it seems logical to assume that the dialect speech was causally related to the spelling difficulty. This finding of dialect interference is congruent with those of studies with other dialects (Kligman et al., 1972; Groff, 1978; and Boiarsky, 1969). Moreover, the finding that a significantly greater percentage of spelling errors made by the Newfoundland fourth graders in a free writing exercise were dialect-based indicates that the interference is not limited to spelling



isolated words in a controlled exercise but that it represents an extra impediment to these students' writing.

Not all the dialect features produced evidence that they were related, at least in a direct way, to spelling difficulty. That is not to deny, however, that the features which appeared not to produce dialect-based errors could have exercised an indirect influence on learning to spell. For example, one could speculate that a dialect speaker, faced with additional instances where the orthography does not consistently represent his speech sounds, could suffer a loss of confidence in the orthographic system such that his mastery in both reading and writing is delayed. In this case misspellings of a more random and unpredictable nature could well result from a general sense of dismay at the task of writing words. It was certainly true that the dialect speakers in this study, the Newfoundland fourth graders, were considerably less adept at spelling than their Nova Scotia counterparts.

The finding that only some features of the dialect interfered with spelling was similar to the results of studies with Black English (Kligman et al., 1972; Groff, 1978). Both of these studies found that the / I / sound preceding / l / as in / still / produced "steal" and "steel" misspellings among both the Black English-speaking second graders and middle school students studied. In the present study seven out of 38 misspellings of the word "still" were of this nature. However, in the case of this Newfoundland dialect the / I / may go either /iy/ or /e/ when followed by /l/. This extra variation showed up as 14 misspellings of "still" as "stell" in the Newfoundland sample. The same pattern was shown by the Newfoundland misspelling of the other words illustrating this feature: "kill" produced one "keill" and nine variations of "kell"; "fill" was misspelled as "feil" once and as "fell" 20 times; and "bill" appeared once as "beil" and 25 times as "bell." Another feature of Black English which Groff (1978) found persisting in the spelling errors of his middle grade sample was the / d / and / t / deletions at the end of past tense forms of verbs. In the present study two items involving this feature produced the largest number of dialect-based misspellings. To a reduced extent the same misspellings were present in the errors of the control group. Kligman et al. (1972) concluded that the / e / sound before nasals as in /ten/ caused difficulty for Black English-speaking second graders and some difficulty for standard English-speaking speakers. However, Groff (1978) found that by the middle grades this difficulty had largely been overcome by his sample. The word "stem" was an item in this study that produced a relatively large number of / e / → / I / misspellings among the Newfoundland sample and a small number among the control group. The same word feature in other phonological environments resulted in only a few predictable errors. The similarity between these selective interference findings may indicate that some features in some environments are intrinsically difficult to perceive and that spelling problems result not simply from the fact of dialect variation but from this variation combined with other factors of difficulty.

Some details of the results of this study may be discussed in terms of theories of the spelling process as they apply to dialect interference. The generate-test theory would predict, in cases where words were unfamiliar as spelling words yet familiar as reading words, that dialect-related mis-

spellings would be more frequent when the error is itself an alternative word. The serial probability theory would, on the other hand, predict errors that are congruent with permissible letter sequences in English. An informal analysis of the 231 dialect-based misspellings produced by the Newfoundland sample showed that 65 per cent of the responses were real words and 35 per cent were nonwords. That is, the majority of errors that the dialect would predict were responses like "bell" for "bill" and "tick" for "thick" as opposed to nonword responses like "beal" and "tek." However, there were some individual words that did not conform to this pattern: "still" was more often spelled as "stell" than "steel" or "steal," and "stem" was realized as "stim" by 14 subjects. Thus nothing more than very tenuous support for the generate-test theory is offered by these incidental findings. Further research might find it useful to pursue the issue of the relative power of the generate-test and the serial probability theories to explain dialect interference in spelling.

The results of this study have implications for the general issue of dialect interference in spelling and for spelling instruction in Newfoundland. From the consistent findings over at least three dialects, the fact of dialect interference in spelling seems well established. However, this is only a rather crude answer to a complex question as the variation among dialect features in their effect on spelling suggests. This variation seems to be both between phonological features and within particular features as inconsistent results between words illustrating the same features are revealed. The research question now should be what features interfere in what circumstances or environments and, from a theoretical perspective, how does this interference occur?

Dialect interference in spelling would seem to be another reason why teachers of language arts should know the features of the local dialect. On Newfoundland's northeast coast, teachers who have a systematic knowledge of the dialect can adopt a more understanding attitude to spelling problems involving past tenses of verbs, front vowel sounds, and consonant cluster reduction. Such teachers would be able to scrutinize spelling texts and workbooks to identify words that contain dialect pitfalls not anticipated by authors from central Canada or the United States. They would be able to adjust their instruction to help students around the friction points between their dialects and more conventional pronunciation as they affect spelling. The necessary knowledge would call for inservice as well as preservice education in dialectology and its implications for language learning, especially for learning to use written language.

#### References

- Boiarsky, C. Consistency of spelling and pronunciation deviations of Appalachian students. *Modern Language Journal*, 1969, 53 (5), 1113-1118.
- Bregelman, F. H. Dialect and the teaching of spelling. *Research in the Teaching of English*, 1970, 4, 129-138.
- Cahen, L. S., Craun, M. J., & Johnson, S. K. Spelling difficulty: A survey of the research. *Review of Educational Research*, 1971, 41 (4), 281-301.
- Chomsky, C. Reading, writing, and phonology. *Harvard Educational Review*, 1970, 40 (2), 287-309.
- Chomsky, N., & Hallé, M. *The sound pattern of English*. New York: Harper and Row, 1968.



- Graham, R. T., & Rudolf, E. H. Dialect and spelling. *Elementary English*, 1970, 47, 363-376.
- Groff, P. Children's spelling of features of Black English. *Research in the Teaching of English*, 1978, 12 (1), 21-28.
- Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudolf, E. H. Phoneme-grapheme correspondences as cues to spelling improvement. Washington: U.S. Government Printing Office, 1966.
- Kligman, D. S., Cronnell, B. A., & Verna, G. B. Black English pronunciation and spelling performance. *Elementary English*, 1972, 49 (8), 1247-53.
- Paddock, H. Preliminary dialect mapping of Newfoundland. In H. Paddock (Ed.), *Languages in Newfoundland and Labrador*. (Preliminary Version.) St. John's: Memorial University Department of Linguistics, 1977, 90-99.
- Peters, M. L. Success in spelling: Study of factors affecting improvement of spelling in the junior school. *Cambridge Monographs on Education No. 4*. Cambridge: Cambridge Institute of Education, 1970.
- Simon, D. P., & Simon, H. A. Alternative uses of phonemic information in spelling. *Review of Educational Research*, 1973, 43 (1), 115-137.
- Tiegs, W., & Clark, W. W. *California Achievement Tests, 1970 Edition*. Monterey, California: CTB/McGraw-Hill, Inc., 1970.
- Trager, G. L., & Smith, H. L., Jr. *An outline of English structure*. Norman, Okla.: Battenburg Press, 1951.
- Venezky, R. L. English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly*, Spring, 1967, 2 (3), 75-105.
- Walker, L. Newfoundland dialect interference in oral reading. *Journal of Reading Behavior*, 1975, 7 (1), 61-78.
- Walker, L., & Paddock, H. Spelling and the Newfoundland dialect. *Morning Watch*, 1975, 2 (4), 9-13.
- Wallach, M. A. Perceptual recognition of approximations to English in relation to spelling achievement. *Journal of Educational Psychology*, 1963, 54 (1), 57-62.
- Wolfram, W., & Whiteman, M. The role of dialect interference in composition. ERIC Document, ED 045 971, 1971.



T. E. KIEREN

*The University of Alberta*  
and

B. SOUTHWELL

*Nepean College of Advanced Education*

## The Development in Children and Adolescents of the Construct of Rational Numbers as Operators

*The purpose of this study was to characterize the growth of the operator construct of rational numbers in children and adolescents. A random sample of 72 subjects from 4 elementary and junior high schools were interviewed while doing 7 direct and 7 inverse tasks in either a "real problem packing machine" or a number pattern setting. There were significant age group performance differences. There was an important age  $\times$  representation interaction effect. Children in the machine setting achieved relatively high scores at age 11-12 which were not duplicated by the pattern group until age 13-14. A nominal analysis revealed 3 phases of rational number performance: a " $\frac{1}{2}$ " phase, a unit operator phase, and a general operator phase. It appears that the machine setting prompted a functional interpretation of equivalence and allowed for successful use of covaried partitioning—these two factors promoting earlier general operator performance. The pattern setting appeared to prompt a more algebraic interpretation of equivalence which was used very successfully by older students, particularly females.*

Fractional or rational number ideas historically have been and continue to be difficult for students in school mathematics. Yet, rational numbers are critically important because they represent a means by which a person can deal numerically with continuous phenomena in the same way counting number ideas can be applied to discrete phenomena.

---

Dr. Kieren completed his doctoral degree at the University of Minnesota. He has been associated with the Department of Secondary Education at the University of Alberta for most of the past 12 years. His research has focussed on the mathematical thinking of children and young adults.

Dr. Beth Southwell completed her doctoral degree at the Institute of Education, University of London, and is currently a Senior Lecturer at Nepean College of Advanced Education, Australia. Her research has centred around the development of mathematical concepts in children and includes investigations in Papua New Guinea.

One reason that the rational number construct proves difficult to learn is its complexity. Kieren (in press) has theoretically posited the existence of five subconstructs which appear necessary for a fully functional fractional number construct. These five are part-whole, ratio, quotient, measure, and operator. Full development of these five constructs by a person should allow him or her to apply these ideas in the wide variety of daily and mathematical circumstances where they are pertinent. Kieren (in press) has also hypothesized the existence of mechanisms which a person uses to build up these constructs. Some of these mechanisms, for example, proportional reasoning, are *developmental* in nature and depend on the onset of various stages (to a certain extent age-related) for their personal usefulness. Noelting (1978a) has carefully documented the characteristics of the stages of such reasoning as it pertains to fractions. Kieren (in press) labels other mechanisms as *constructive*. He theorizes the existence of mechanisms for developing rational number constructs which would parallel the central role played by counting in whole number development.

The study reported here focusses on the development of the operator subconstruct of rational numbers. It represents an extension of a previous study of rational numbers as operators (Kieren & Nelson, 1978) in that it attempts to study more clearly and carefully phases in development and mechanisms involved.

The operator construct is an important focus of study for two reasons. First fractional or rational numbers are by nature two-sided; they are simultaneously additive vectors and functions which can be composed. The operator subconstruct is the clearest representation of rational numbers as numbers to be multiplicatively composed. The second reason for their importance is curricular. Approaching the study of fractions through operators has been suggested by Dienes (1971) and widely implemented in the school curriculums of the Federal Republic of Germany (Padberg, 1978). Like many other approaches to the teaching of fractions, it is important to see how and why it might be made most effective in terms of the learner's construct development.

There are two major purposes of this study. The first is to trace the development of the operator construct. Does the quality of this construct change with age and, in particular, are there quantitatively identifiable growth points or levels as tentatively seen in a previous study (Kieren & Nelson, 1978)? How do developmental and constructive mechanisms, if observable, affect this growth?

The second major purpose deals with the representation of the construct. In Kieren and Nelson (1978), the fractional operator was represented by a machine, the nature of which the young person was to discover. This representation allowed the person to manipulate objects and see the idea in a "real" problem setting. Such settings prove difficult to include in curricular materials. A simpler curricular approach to operators is found in the study of patterns in pairs of numbers. An important question is, "How does the pattern representation as opposed to the machine representation affect the way in which the young person thinks about the fractional number?"

There are three minor purposes of the study. In a previous study Kieren



and Nelson (1978) found that the notion of “1/2” formed a stage or level in fractional thinking in the operator setting. This is confirmed by the work of Piaget, Inhelder and Syzminska (1966) and Noelting (1978a) in other ways. Kieren and Nelson further observed that older children and adolescents persisted in making “1/2 errors” in situations they found unfamiliar. Because in the previous study the first task was a “1/2 task” it was thought that possibly this task induced a “set” which crept into later tasks. In the current study, subjects were randomly assigned to one of two task orderings, one which started with a “1/2 task” and one which did not.

A second minor purpose was to further establish the fact that a subject who could handle the direct fractional operator could also do its inverse (Kieren & Nelson, 1978). This result is seen as a particular manifestation of the mechanism of reversibility and should be well established in the subjects of ages used in this study.

The third minor purpose was to look at the notion of composition of functions. In other mathematics educational research (Lovell, 1971; Thomas, 1975) composition of function is seen as an advanced notion. This present study allowed one to consider composition in the setting of fractional operators and is indeed central to this construct. (A detailed report of this aspect of the study can be found in Kieren and Southwell, 1978).

Procedures

Sample

As in the Kieren and Nelson (1978) study, the sample for this study was chosen from students in Grades 4 to 8. This was done because these are the levels of most significant instruction in rational numbers. In addition, students at these ages should have reached the developmental level in which they can exhibit reversibility and have primitive proportional reasoning skills (Noelting, 1978a). Thus, 72 subjects were selected randomly from the populations of two elementary and two junior high schools. There were 16 subjects at each of Grades 4, 5, and 6 and 12 subjects at Grades 7 and 8. One elementary and one junior high school were from the Public schools while the other two were from the Catholic school system of a large western Canadian city. The schools were in either inner city or suburban locations. For the purpose of analysis this sample of 72 was separated by age into five categories as listed in Table 1.

To test the appropriateness of the lower bound of the age range, four children from Grade 2, aged 7 and 8, also were tested. In addition, 4 high achieving Grade 10 students, aged 16, were tested. In the previous study there was only one person, aged 13, who reached the highest performance category. The four Grade 10 students were tested to see if they would perform uniformly at a high level and thus further validate the task set.

TABLE 1  
AGE GROUP SIZES

Age in years	< 10	$10 \leq a < 11$	$11 \leq a < 12$	$12 \leq a < 13$	$a \geq 13$
<i>n</i>	11	20	13	13	15



The testing was done during April and May of 1978 in a manner documented below by the two experimenters using a well-defined protocol. Each child was tested separately in a room provided by each of the schools. All of the children in the sample had received some instruction on fractions or rational numbers and this was uniform, at least in terms of objectives, across schools at a grade level. In all cases, the subjects were drawn from more than one classroom in a school.

Upon coming to the testing area, students were randomly assigned to either a machine or pattern setting and, within settings, assigned to either a “1/2 first” or “1/3 first” task sequence. These procedures resulted in the categories outlined in Table 2. The two experimenters randomly administered the machine and pattern settings.

TABLE 2  
NUMBERS OF SUBJECTS ASSIGNED TO TESTING SETTINGS

	Machine	Pattern
1/2 first	19	17
1/3 first	15	21

Instruments

*Machine.* A series of fractional tasks was presented using two different modes. The first mode involved the use of a coloured box with input and output doors. The subjects were told that the box represented a packing machine which was programmed to behave in a certain way. As objects, represented by small pieces of cardboard, are fed into the machine, they are “packed” into larger packages, represented by a larger piece of coloured cardboard.

The subjects were asked to feed in a certain number of objects and count the resulting output. After up to six trials of this kind, they were asked to predict what the outcome would be on five specific tasks presented to them on a sheet with IN and OUT columns. Upon completion of the sheet, each subject was asked to explain how the machine was working. A written record of free responses was kept.

A second sheet with IN and OUT columns was handed to each subject for completion. “What would I put into the machine to get 7 out?” was the form of question asked. Upon completion of the five tasks on this second sheet, the subject was again asked to explain how the machine was working and his free responses were recorded.

The subject was then told that the machine had been reprogrammed and a similar sequence of questions and instructions were given using a different fractional operator. In all, the subject was presented with 7 sets of 5 direct and 5 inverse items, representing a total of 14 tasks. Variables used in these tasks were (i) whether the operator involved was a unit ( $1/2$ ,  $1/3$ ,  $1/3 \times 1/2$ ) or a non unit fraction ( $3/4$ ,  $2/3$ ,  $3/4 \times 1/2$ ,  $3/4 \times 2/5$ ) and (ii) whether the operator was single ( $1/2$ ,  $1/3$ ,  $2/3$ ,  $3/4$ ) or composite ( $1/3 + 1/2$ ,  $3/4 + 1/2$ ,  $3/4 + 2/5$ ). (See Kieren and Nelson, 1978, for details of this procedure.)

In the latter case, two machines were used. The first, with which the subjects were already familiar, put objects into packages. The output of this first machine became the input for the second machine which packed the packages into cartons (represented by larger pieces of cardboard). For instance, for the composition,  $3/4 \times 1/2$ , sixteen objects were fed into the first ( $3/4$ ) machine and twelve packages came out. These twelve packages were fed into the second machine and six cartons resulted. As in the former setting with one machine, the subjects were given up to six trials and were then asked to predict the outcome in five direct and five inverse tasks on this composite setting.

A third variable was whether subjects were presented with a sequence of tasks beginning with “ $1/2$ ” or whether they were presented with a sequence beginning with “ $1/3$ ”. In the former case, the sequence was  $1/2, 2/3, 1/3, 3/4, 1/3 \times 1/2, 3/4 \times 1/2, 3/4 \times 2/5$ . The second sequence was  $1/3, 3/4, 1/2, 2/3, 1/3 \times 1/2, 3/4 \times 1/2, 3/4 \times 2/5$ .

The sets of tasks were presented to the subjects individually. The subjects were given as much time as they needed to complete as many of the tasks as possible. Each subject followed one of the above sequences and proceeded until he or she was unable to make any predictions or explain how the machine(s) was functioning. Predictions were recorded by each subject on the appropriate sheets for the 14 (7 direct, 7 inverse) tasks. In addition, the experimenter kept a written record of subjects’ discussion and descriptions of the machine.

*Pattern.* The second mode involved the use of patterns displayed on pieces of cardboard, with numerals linked together by an arrow, e.g.

$$12 \rightarrow 4$$

Up to six cards with the same pattern (e.g.,  $1/3$ ) were displayed and the subject asked to find the pattern. This was done in the same way as with the machine mode. After several cards were displayed, the subject was asked to predict the outcome using the same pattern; for example

given	$12 \rightarrow 4,$
	$9 \rightarrow 3,$
	$27 \rightarrow 9,$
	$60 \rightarrow 20,$
and	$120 \rightarrow 40,$
	$6 \rightarrow ? \text{ etc.}$

The subject recorded his or her predictions on five tasks on sheets as with the machine mode. For inverse items, the subject was asked what would have to be on the left hand side of the arrow to give  $n$  on the right hand side.

The tasks presented to the subjects using the pattern mode were the same as those given using the machine. Thus, the same three variables were present in this setting. The subjects recorded their predictions individually and the experimenter kept a written record of discussion and descriptions used by the subjects. These test items were such that the required result was always a whole number. In addition, the items were such that three items tested the operator on “small” numbers ( $n \leq 40$ ), one on numbers from 50 to 200, and one item on numbers in the thousands. This was done to assess whether the subject was using a single strategy for each operator and



whether this strategy was personally applicable across a wide domain of numbers. Thus for small numbers a subject might see a coincidental but successful pattern which would not carry over to larger numbers. Each subject was given as much time as was required to complete the tasks and proceeded until unable to go any further. In analysis of results, mastery was assessed on success of 4 of 5 items, at least one of which would have to incorporate a "large" number.

Both the machine and the pattern mode proved to be viable settings. All subjects tested were able to function within these settings. The pattern mode, however, is considered to present a more abstract task than the machine in which the subjects were able to manipulate concrete objects providing both motivation and clarification of the task. The time for each interview was dependent on the needs of the subject; actual times ranged from 13 to 70 minutes.

All students did not complete all settings. For example, subjects who were unable to do  $3/4$  tasks were not asked to attempt  $3/4 \times 1/2$  or  $3/4 \times 2/5$ . The minimum number of tasks attempted was one and the median number, 8 of 14.

### *Questions and Hypotheses*

In an important sense this research was hypothesis-generating. It allowed for a detailed consideration of child and adolescent behaviour against controlled protocols with two representations of the operator construct of fractional or rational numbers. From the records of subject behaviour, one can attempt to develop a reasoned explanation or set of hypotheses which describe the pattern of behaviour. The following hypotheses assisted in the development of such an explanation.

—There is no difference in mean performance by students faced with "1/2 task first" sequence as opposed to the "1/3 task first" sequence.

—Performance on inverse tasks is independent of performance on direct operator tasks.

—There exist differences in mean performances of groups at different age levels with older subjects performing at a higher level.

—There are no differences in mean performance by subjects in the machine and pattern settings.

—There are stable categories of performance which typify the development of the operator construct.

The data used in testing these hypotheses and in generating the theoretical discussion below come from two sources. The performance of subjects on each of 14 tasks (5 items per task) was recorded. The split halves reliability of results on these items is .91. The transcripts of subjects' verbal and nonverbal reactions to each task faced were also available for study.

## *Results*

### *Preliminary*

Before turning to the major purposes of the study, two preliminary results are of interest. The first deals with the testing sequence. Because "1/2" holds such an important position in rational number construct development,



the “set” generated by presenting the “1/2 task” first may have affected the results of the study. Thus the mean performances of the “1/2 first” vs. “1/3 first” were compared at each of the age levels of the study. With the exception of youngest age level ( $a < 10$ ) where the “1/2 first” group was higher, ( $t = 2.77$ , sig. at .02) there were no significant differences in the mean performances of the two groups at any age level. The results in the youngest age group were greatly affected by the performances of one person in each sequence category. This and a study of 1/2 errors in subject protocols leads to the conclusion that, as will be seen below, although “1/2” is a stage in development, the sequence “1/2 first” does not account for differing performances or errors among students.

The second preliminary result deals with performance on direct and inverse items. Table 3 shows the number of subjects mastering (4 of 5 items correct) related direct and inverse tasks (for example, numbers of subjects mastering both “1/2 direct” and “1/2 inverse” items). Thus, although the inverse items presented somewhat more difficulty than direct items (42.5% vs. 48.5% mastery performances), there did exist a condition of dependency. Therefore a subject who mastered a direct task likely would also master the related inverse task. It appears that it is the nature of the operator (1/2, 3/4, etc.) which determines performance rather than the aspect of inverse within an operator.

TABLE 3  
NUMBERS OF PERFORMANCES IN VARIOUS MASTERY CATEGORIES  
WITH RESPECT TO PAIRED DIRECT AND INVERSE TASKS

		Inverse		
		Mastered	Non-mastered	
Direct	Mastered	216	42	258 (48.5%)
	Non-mastered	10	264	274 (51.5%)
		226 (42.5%)	306 (57.5%)	532

$p(\chi_2 = 351) < .01$

A final preliminary result relates to the performance of Grade 2 students (ages 7, 3 to 8, 5). Only one of 4 students was able to do any of the tasks, a girl aged 7 years 3 months, and she was able to cope only minimally with the 1/2 task. This result is as hypothesized and is consistent with the findings of Noelting (1978a). He found that the onset of the first level of the equivalence class reasoning stage was at age 8 and the more general equivalence class stage at age 10. Thus the younger students would, at best, be able to cope with the “1/2” tasks of this study.

Central Results

After up to six trials with feedback on each of the task settings, the subject was asked to predict, in writing, the output (or input) of a machine or pattern for a given task. There were 5 direct and 5 inverse items per task.

TABLE 4  
MEAN PERFORMANCES OF SUBJECTS IN VARIOUS AGE  
AND REPRESENTATIONAL CATEGORIES

Mode	Ages in Years					Total
	a < 10	10 ≤ a < 11	11 ≤ a < 12	12 ≤ a < 13	a ≥ 13	
Pattern	11.4	26.1	37.8	32.3	48.3	31.2
Machine	14	24.8	43.2	42.6	40.9	32.9
Total	12.6	25.5	40.7	36.3	44.4	

TABLE 5  
ANALYSIS OF VARIANCE OF ACHIEVEMENT SCORES USING  
REPRESENTATION AND AGE AS BLOCKING VARIABLES

Source	Sum of Squares	df	Mean Square	F
Representations	213	1	213	.88
Age	7233	4	1808	7.47*
Interaction	3902	4	975	4.03*
Within Cells	15034	62	242	
Total	26382	71		

\* Significant at a .01 level.

Thus there was a possible total of 70 correct predictions. Table 4 presents the mean achievement of each age level group as well as for each age level group in each of the representational settings.

Table 5 reports the analysis of variance performed using representation and age as variables.

It is clear from this analysis that there is a significant age effect as would be expected. While there appears to be no general representation effect, this result is tempered by the interaction which can be vividly seen in Figure 1. From this figure and from related contrast analyses, the following conclusion can be drawn. For the machine groups there are significant achievement gains from age (<10) group to age (10 ≤ a < 11) group and from that group to the age (11 ≤ a < 12) group, the latter a dramatic change. There is no significant growth seen at age (12 ≤ a < 13) or age (a ≥ 13) groups. With the pattern groups significant gains appear from age (<10) to age (10 ≤ a < 11), from age (10 ≤ a < 11) to age (11 ≤ a < 12) and a dramatic gain from age (12 ≤ a < 13) to age (a ≥ 13). Thus, it appears that the machine group reaches a high level of performance at age 11-12 while with the pattern group, high level of performance comes at a later age (a ≥ 13).

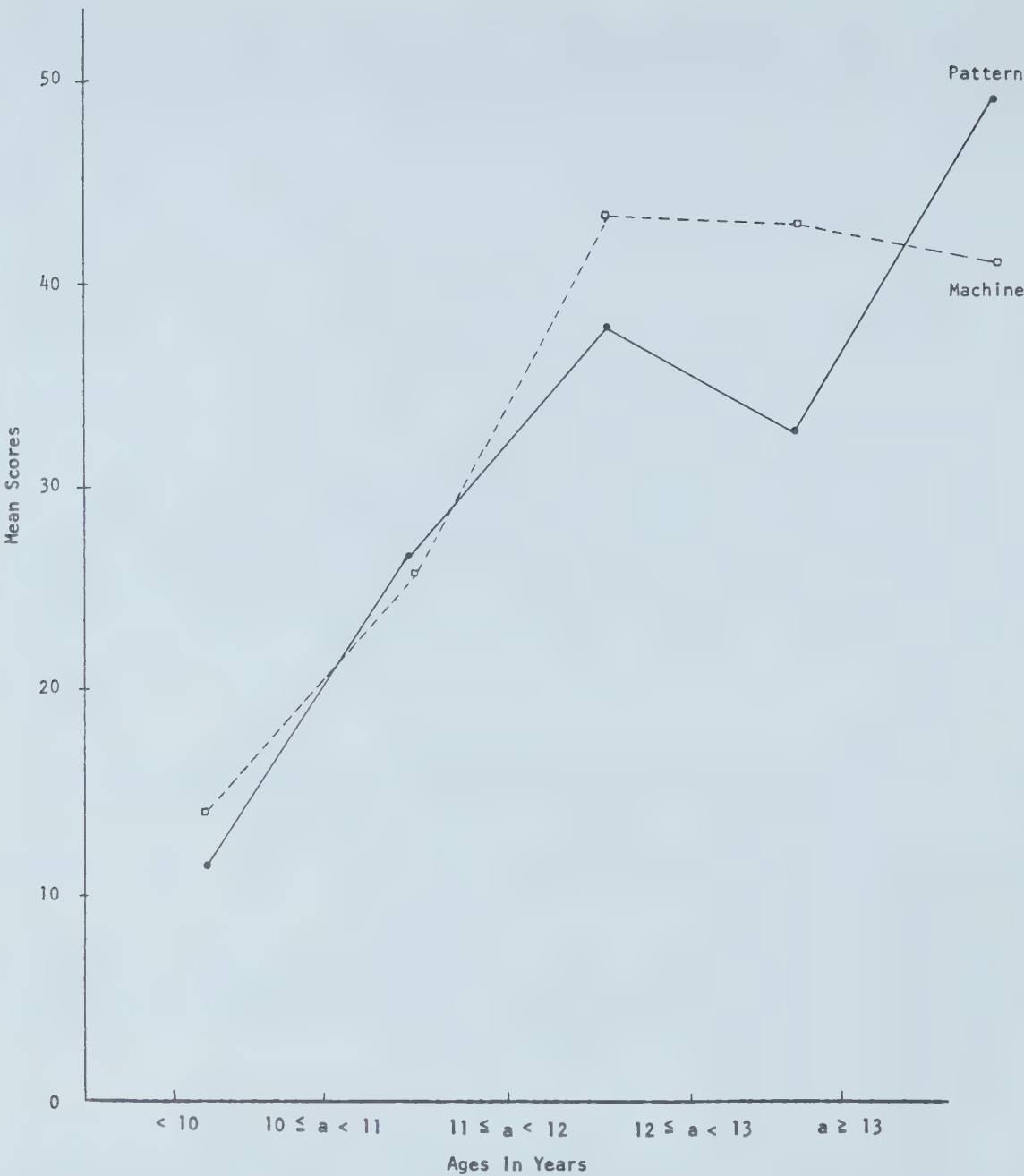


Figure 1. Performances by age and representation.

*Nominal Analysis*

The task performances and explanations given by children and young adults in each of the settings were analyzed to determine if certain “stages” in the development of the construct might be observed. From this analysis seven levels or stages appeared.

- I: Mastered none of the 14 tasks.
- II: Mastered the “1/2” tasks (direct or inverse) only.
- III: Mastered “1/2” and “1/3” tasks.
- IV: Mastered unit tasks including composition.
- V: Mastered some non-unit tasks (2/3, 3/4 compositions).



VI: Mastered all tasks except  $\frac{3}{4} \times \frac{2}{5}$ .

VII: Mastered all tasks.

A purpose of this study was to search for performance categories. Hence no formal attachment of criterion tasks to stages was made a priori and the corresponding hierarchy analysis was not formally considered. It is interesting to note that well over 90% of subjects in level  $n$  performed all of the tasks associated with level  $n - 1$ . Conversely less than 10% of the subjects who failed a level  $n$  task mastered any task at a level above level  $n$ . For example all subjects who mastered  $\frac{3}{4}$  also mastered  $\frac{1}{2} \times \frac{1}{3}$ , and of those failing " $\frac{1}{2} \times \frac{1}{3}$ ," only 1 subject mastered a non-unit task. Thus, though not formally tested, the levels appear to be hierarchical in nature.

TABLE 6  
LEVELS IN THE GROWTH OF THE OPERATOR CONSTRUCT

	< 10		10 ≤ a < 11		11 ≤ a < 12		12 ≤ a < 13		a ≥ 13		Age Median	Age of Accession	Number of		Total Both
													Girls	Boys	
VII			GP			BM BM			GP GP		12+	> 13	3	2	5
VI				BM		BM	GP	BM	GM		12+	> 13	2	3	5
V				BM		BP BP BP		BM BM	GM GP GP	BM BM BP	13	> 13	3	9	12
IV			GM	BP BP BP BP	GM	BM	GP GP GM	BP BP BP BP	GM GM GP GM GP	BM	12+	11+	10	10	20
III		BM BP	GM GP	BP	GP GP		GP				10+	10+	5	3	8
II	GM GP	BM BP	GM GM GM	BM BP	GM GP		GM				10, 6	< 10	8	4	12
I	GM GP	BP BP BM	GM GP	BP BM		BM					10 10	< 10	4	6	10

Table 6 presents a record of the numbers of subjects classified at the various levels. Subjects are identified by representation faced (M or P) and by sex (G or B). The two experimenters studied each subject's record and classified subjects independently; there was 100% agreement on the classification of subjects. The table also gives the median age group for each level and the age of accession—the age group at which over 50% of the group are at that particular stage or above. For example, the age of accession for level IV is  $11 \leq a < 12$ , because for that group (and those older) 50% are at or above level IV, while less than 50% of the next younger group is at level IV.

Table 6 confirms the general pattern of performance gain with age discussed earlier. Yet at each age level, particularly the middle three, there is a wide distribution of performances. This is of particular interest because it is at these age levels that the great preponderance of fractional number instruction takes place. It is also interesting to note that in each performance category, the ages of the subjects vary considerably. There is

one girl aged 10 who mastered all tasks and one boy aged 11 who mastered none. Though this distribution occurs, it is useful to note that over 90% of subjects aged 12 and over fell in levels IV and above, while 74% of children below age 11 were classed below level IV.

The performance of the high achieving grade 10 students lends further validity and credence to the tasks and the category system. All of these subjects (their performances have not appeared in any previous analysis) were classified at level VII and their mean performance score was 66.5.

Further consideration of the nominal analysis above suggests that there are three general phases of performance or operator constructs. The first of these represents a rather primitive fractional construct and can be associated with "1/2 only" performance. Vestiges of this construct are seen in the "1/2 errors" which continue to be made on advanced tasks by subjects up to the age of 13 in this sample. The second operator construct is that of the unit operator and represents control over unit operators in general, including composition or multiplication of such operators. It was found that over 90% of the students who mastered " $1/3 \times 1/2$ " also could name it with the usual response being "divide by 6". The third phase of the operator construct is a general operator notion which incorporates unit and non-unit operators and compositions.

Thus Level II in Table 6 represents the "1/2 operator." Level III is transitional and leads to the "unit operator construct" represented by Level IV. Level V is a primitive general operator transition phase leading to the "general operator" construct. Level VII represents performance which is based on abstract manipulation of operators and may represent a fully general construct.

*Sex Differences*

It was not anticipated that there would be sex differences in performance on operator tasks. Thus, sex was not considered a sampling variable. Nevertheless the random sample contained 35 females and 37 males. The means for males was 34.3 and females 28.9; this difference was not significant ( $t = 1.05$ ).

However, perusal of Table 6 suggests that 10 of 14 general operator phase performances by boys occurred under the machine representation while 6 of 8 general operator performances by girls occurred under the pattern representation. This sex by representation "interaction" is seen in Table 7. The sex difference under the machine representation is significant ( $t_{32} = 2.98$ ) at the .01 level while the difference under the pattern mode is not ( $t_{36} = 1.11$ ).

TABLE 7  
MEAN SCORES BY SEX AND REPRESENTATION

	<u>Machine</u>	<u>Pattern</u>
<u>Male</u>	42.76 ( $n = 17$ )	27.15 ( $n = 20$ )
<u>Female</u>	24.41 ( $n = 17$ )	33.78 ( $n = 18$ )



*Discussion*

How or in what patterns do young persons think of fractions or rational numbers as operators? Both the analysis of mean performances and the nominal analysis suggest a pattern of growth over time. There appear to be three major phases in the operator construct, the "1/2 operator" phase, the unit operator phase, and the general operator phase. With transitions, the nominal analysis suggested 7 levels of performance across the ages 9-14.

What mechanisms does a young person use in coming to grips with rationals as operators? Noelting (1978b) has clearly shown several relevant phases in proportional thought. Particularly Noelting's phase IIa (p. 9) involving the equivalence class of the 1 to 1 ratio is sufficient to account for behaviour with "1/2 operators." His age of accession to this stage (8, 1) seems to relate well to age of accession to the 1/2 operator construct.

How does a young person gain a broader construct of fractional operators. While "equivalence class" thinking is germane to all 14 tasks in this study, it would appear that the mechanism of "partitioning" plays a central role as well. Partitioning is defined here to mean the division of a set into subsets of the same size and manifests itself in subjects' "division" statements. Thus a subject correctly pairs 2 with 6 in a "1/3 task" and explains "divides by 3." This kind of response was prevalent in this sample beyond age 10, with younger students more frequently giving subtractive explanations for their predictions.

The passage from the unit operator level to the general operator phase involves a more complex use of the mechanism of partitioning. For example a successful subject (age 13) pairs 90 with 60 in the 2/3 task and says "divide 90 by 3 and then multiply by 2" recognizing that whatever divides the input element 3 times divides the output element twice. An unsuccessful subject (age 10) realizes partitioning is involved but manipulates her explanation to fit her erroneous answers:

$$90 \rightarrow 50$$

"Three goes into 90 and 5 goes into 50."

It appears that control of the general fractional operators requires the partitioning of two sets or numbers but also coordinating this partitioning or, as Noelting (1978b, p. 15) suggests, using covariational thinking.

This covaried partitioning strategy occurred mainly in the subjects under the machine representation. In the pattern representation, a pattern explanation frequently accompanied a correct response. Thus in pairing 24 with 16 in the 2/3 task the subject would say "Well I knew 12 went to 8 so I just doubled it and got 16."

In some respects, seeing the pattern, particularly in non-unit fractional operators, represents more sophisticated thinking than the more rule-oriented covaried partitioning. This might explain the fact that the machine group reached a higher level of performance at a younger age than did the pattern group. Yet this use of partitioning also led to some clumsy if correct strategies as exemplified by this explanation "Well you divide 100 by 4 and then subtract that from 100." The high achieving grade 10 student who gave such an explanation also found great initial difficulty in handling non-unit inverse tasks. She never did reveal her final procedure but would say "I



would just think, what number do I divide by 4 and then subtract to get 60?" Although this strategy is plausible, her final response speed indicated some other procedure was in effect.

It would appear that students tested under the two representations used different strategies particularly with respect to non-unit operators. The machine representation appeared to induce a focus on a rule relating input to output, this rule most often involving covaried partitioning. The pattern representation was best handled by a recourse to known results; thus  $24 \rightarrow 18$  is seen as  $3(8 \rightarrow 6)$  and seemed to require knowledge, general proportional reasoning or, as Noelting sees it, "Common denominator thinking." It would be of significant theoretical interest to administer the operator tasks and the Noelting test to the same subjects and test the above conjecture. It would appear that young persons can cope with operator settings, at least in the machine mode, by "substituting" the mechanism of partitioning for a more sophisticated control over ratios.

An alternative explanation of the mode by age interaction concerns the notion of equivalence class. A subject who says "divide by 4 and subtract" sees the various " $3/4$ " items as being determined by a single rule. In the case of a subject who sees  $(8 \rightarrow 6)$  as  $1/2(16 \rightarrow 12)$ , the idea of equivalence as a class of ordered pairs which fit together appears to prevail. The former explanation of an item simply involves "working it out" while the latter appears to involve at least an implicit idea of a set of ordered pairs, the latter being a more formal mathematical notion as well as a more algebraic notion. It might be hypothesized that the machine mode sponsors the application of partitioning which in turn permits the use of a more primitive notion of equivalence in dealing with fractions as operators.

Nelson (in press) has suggested that "real" problems allow children to exhibit behaviours which seem beyond their predicted level of thought. Thus the "machine" representation might be an appropriate introduction to the operator construct with the less playful pattern representations and standard symbolic representations coming quite a bit later. To an extent this introduction to fractions obtains in various German curriculums today for students of about age 11. This age corresponds well with the results of Noelting (1978a) and this study. It would appear that a useful prerequisite for the study of general operators would be many different experiences with partitioning, particularly of discrete sets, as well as extensive work with unit fractional operators.

Before acting on the above curricular suggestions several questions need answering. From this study and to a much greater extent from the results of the previous study (Nelson & Kieren, 1978), it would appear that work with general rational operators would best be reserved for the junior high school. Is this the case, or can non-unit rationals as operators be successfully taught earlier?

It would appear that representation affects both attainment of the operator construct and the mechanisms used by young persons in building up the construct. To what extent is this true? What impact does development in the Noelting sense and instructional representation have on this result? Finally within this context, is the sex by representation interaction generally supportable? If so, what are the curricular impacts of this effect?

There appear to be three phases in the growth of the fractional operator construct. Moreover, there appear to be discernible developmental and constructive mechanisms used by young persons in building up this idea. The relationships of the operator subconstruct to other rational number subconstructs need to be illuminated and relevant learning experiences for young persons aged 8 to 15 devised.

*References*

- Dienes, Z. P. *The elements of mathematics*, New York: Gordon and Bream, 1971.
- Kieren, T. E. The rational number construct—Its elements and mechanisms. In T. E. Kieren (Ed.), *Recent research on number learning*. Columbus: ERIC/SMEAC, in press.
- Kieren, T. E., & Nelson, L. D. The operator construct of rational numbers in childhood and adolescence—An exploratory study. *The Alberta Journal of Educational Research*, 1978, 24, 22-30.
- Kieren, T. E., & Southwell, B. A comment on function composition. Edmonton: University of Alberta, 1978.
- Lovell, K. Some aspects of the growth of the concept of a function. In M. Roszkopf, L. Steffe, & S. Taback (Eds.), *Projection cognitive development research and mathematical education*. Reston, VA: NCTM, 1971, 12-33.
- Nelson, L. D. A hopeful trend in problem solving research. In R. Lesh (Ed.), *Applied problem solving*. Columbus: ERIC/SMEAC, in press.
- Noelting, G. *La construction de la notion de proportion chez l'enfant et l'adolescent et les mécanismes d'équilibration*. Québec: Ecole de Psychologie, Université Laval, 1978. (a)
- Noelting, G. Constructivism as a model for cognitive development and (eventually) learning. Québec: Ecole de Psychologie, Université Laval, 1978. (b)
- Padberg, F. *Didaktik der Bruchrechnung*, Freiburg: Herder, 1978.
- Piaget, J., Inhelder, B., & Syzminska, A. *The child's conception of geometry*. New York: Basic Books, 1966.
- Thomas, H. L. The concept of function. In M. Roszkopf (Ed.), *Children's mathematical concepts*. New York: Teacher's College Press, Columbia University, 1975, 145-172.

B. BARBARA ADAMS

and

LLOYD O. OLLILA

*The University of Victoria*

## The Relationships of Language Concepts, Auditory Comprehension, Visual Perception, and Spatial Relations as Predictors to Reading Achievement in First Grade

*The investigators developed indices to measure some aspects of auditory comprehension, visual perception, and basic language concepts at the reading readiness level in order that new knowledge might be added to the construction of tests which are more accurate predictors of reading achievement in the first grade. Multiple linear regression analysis was performed to determine which subtests of the reading readiness tests were the best predictors of reading achievement. Results show that the Learning Rate subtest of the Canadian Readiness Test and Spatial Relations Concepts of the Experimental Reading Readiness Test were highly significant predictors of reading achievement on all three subtests of the Bond-Balow-Hoyt. The sex differences in reading readiness were also investigated. The Experimental Reading Readiness Test indicated no significant differences between boys and girls in reading readiness levels. On the Canadian Readiness Test, however, the girls were superior on two of the subtests, Word Matching and Semantics. In reading achievement on the Bond-Balow-Hoyt: The New Developmental Reading Tests, there were no significant differences between first grade boys and girls.*

Some factors which educators feel contribute to reading readiness are auditory comprehension, auditory discrimination, richness of verbal concepts, general mental ability, knowledge of numerical and qualitative

---

Barbara Adams is an Educational Psychologist specializing in the field of Learning Disabilities and Reading Problems. She received her Master of Arts Education from the University of Victoria and is presently employed by the Greater Victoria School District.

Lloyd O. Ollila is Associate Professor of communications and social foundations in education at the University of Victoria, Canada. He received his B.A. and Ph.D. degrees at the University of Minnesota. His current research is in the area of children's word recognition, relationships between writing and reading, and reading readiness skills.



relationships, the sensory-motor abilities required in hand writing, the ability to attend, and the ability to follow directions (Johnson, 1969).

Standish (1959) and MacGinitie (1969) concluded from their review of the literature that the best predictors tend to be those tasks that are the most similar to the criterion, in other words a test of reading achievement. MacGinitie (1969) gave the following reasons for the predictive validity of tasks that resemble reading. First, if the predictor task is similar to reading, abilities required for the success in reading are measured by the predictor, even if it is not clear just what those abilities are. Second, predictor tasks measure both the environmental and motivational factors which are likely to encourage a child to become a better reader. Parental interest and encouragement will foster continued progress in the development of reading skills. Also the child who is curious enough to inquire about printed words will most likely continue to do so.

Many studies have shown that letter names are the best predictors of reading achievement. In order to name a letter correctly, a child must be able to discriminate visually one letter shape from another. Gates and Bond (1936), Barrett (1965), and de Hirsch, Jansky and Langford (1966) found that visual discrimination of words and letters was one of the best predictors of reading success.

The relative value of a child knowing the letter names, however, has been questioned by some researchers. Samuels (1971) and Ohnmacht (1969) argued against the importance of letter naming and claimed that letter sound was the more relevant variable. Ohnmacht demonstrated that children trained in letter sound performed significantly better than those trained in letter name. Many educators have jumped to the conclusion that because letter name knowledge is one of the best predictors of reading success that this should be taught before teaching the child to read. Barrett (1965) suggested that the ability to recognize the letters of the alphabet may be partially a reflection of past experiences with a variety of written materials which enable these children to recognize letters. These experiences may also help these children become better readers later. Barrett warned that one should not infer from his study that teaching children to recognize letters by name will necessarily ensure success in beginning reading.

These investigators are of the same opinion as MacGinitie (1969), Spache & Spache (1971), and Weintraub (1967) who concluded that test instruments must be found that are better predictors of reading achievement than those now in existence. The first grade teacher must know which children are ready for the instructional methods and/or the materials which she will be using in her classroom. For the children who are not ready, she must know their individual needs so that special readiness programs and materials can be prepared. Since it is not economically feasible to test all first graders individually, we must find group tests that are better predictors of reading readiness and are more useful as diagnostic tools for instruction. Research on reading readiness frequently fails to deal adequately with methods and materials for which children are required to be ready (MacGinitie, 1969). Goodacre (1971) stated that in the past readiness tests have been crude measures only, and therefore, identify only the extremes. In other words they could show only who was reading and who was not

reading. She also concludes that what happens to those children who have not started to read will depend on the instruction they receive from their teacher and her ability to diagnose the children's instructional needs.

Considerable re-examination of measures of reading as predictors of reading achievement is needed. Attempts should be made to define and measure children's developmental factors and to isolate and describe the functions in the reading process. Reading is a process in which the child's developmental characteristics assume varying importance at different stages.

### *Null Hypotheses*

To investigate the remaining purposes of this study, the following null hypotheses were formulated:

1. On the investigator's *Experimental Reading Readiness* subtests and on the *Canadian Readiness* subtests, there will be no significant difference for sex.
2. On the *Bond-Balow-Hoyt: The New Developmental Reading Tests* (Bond, Balow & Hoyt, 1965), there will be no significant difference for sex.
3. On the *Experimental Reading Readiness Test*, the *Canadian Readiness Test* (Evanechko, Ollila, Downing & Braun, 1970), the *Bond-Balow-Hoyt: The New Developmental Reading Tests*, there will be no significant correlation between the various subtests.
4. On the *Experimental Reading Readiness* subtests and the *Canadian Readiness* subtests, there will be no significant predictors of reading achievement on the three subtests of the *Bond-Balow-Hoyt: The New Developmental Reading Tests*.

### *The Purpose of the Study*

The primary purpose of this study was to develop new indices to measure the level of reading readiness of beginning first graders. It was not the authors' purpose to duplicate the reading readiness subtests that are already found to be good predictors of reading achievement. One area of auditory comprehension, visual perception, spatial relations, and language were considered. The specific aims were:

1. to identify and develop indices of reading readiness in the following areas: (a) Predicting Outcomes, (b) Basic Language Concepts, (c) Spatial Relations Concepts, and (d) Auditory Reception.
2. to determine which are the best reading readiness predictors of reading achievement,
3. to determine if there are significant differences in reading readiness and reading achievement scores between boys and girls on the tests used in this study.

### *Method*

#### *The Sample*

Subjects used in this study were one hundred and fifteen children from seven first grade classrooms randomly selected from schools located in the Sooke School District on the lower Vancouver Island in British Columbia.



Sixty-two boys and fifty-three girls from rural and urban areas were involved in the investigation.

#### *Contents of the Experimental Reading Readiness Test*

The four following subtests were developed and administered to the first graders in September, 1975:

Test One: *Basic Language Concepts*—This 30-item test is designed to measure the child's basic language concepts or understanding of words which he has acquired through experience.

Test Two: *Predicting Outcomes*—This 30-item test is designed to measure the child's auditory comprehension and his ability to predict what will happen next when the context is given orally by the examiner.

Test Three: *Spatial Relations Concepts*—This 30-item test is designed to measure the child's concepts of space and his experiences and understanding of words, such as over, under, beside, and biggest.

Test Four: *Auditory Reception*—This 30-item test is designed to measure the child's ability to derive meaning from verbally presented statements and the ability to respond with a simple nonverbal *yes* or *no* by circling the word *yes* and the *happy face* if the statement is correct or the word *no* and the *sad face* if the statement is not correct.

#### *Contents of the Canadian Readiness Test*

Test One: *Technical Language of Literacy*—This 16-item test is designed to measure the child's knowledge of technical terms used in describing language, such as, letter, word, and numeral.

Test Two: *Letter Recognition*—This is a visual discrimination test of 21 items. It is designed to measure the child's ability to recognize both upper and lower case letters.

Test Three: *Word Matching*—This is a 21-item visual discrimination test.

Test Four: *Beginning Sounds*—This 19-item test is designed to measure the child's ability to discriminate between sounds at the beginning of words.

Test Five: *Semantics*—This 12-item test requires the child to classify stimuli by circling the three pictures, from a group of five, that belong together.

Test Six: *Learning Rate*—This subtest of 22 items is designed to measure the child's ability to learn sight words. After the teacher teaches ten sight words, the child is required to select from a row of three words one of the specific words named by the teacher.

#### *Contents of the Bond-Balow-Hoyt:*

##### *The New Developmental Reading Tests*

Part I: *Word Recognition*—This is a 36-item four-choice test designed to assess the pupil's word-recognition ability. The pupil is required to read the four words and mark the one that matches the picture.

Part II: *Comprehending Significant Ideas*—This 40-item test is designed to measure the child's ability to comprehend ideas expressed in a short paragraph.



Part III: *Comprehending Specific Instructions*—This 32-item test is designed to measure the child’s ability to comprehend specific instructions which become more difficult as the child progresses through the test.

*Administration of the Tests*

The authors’ *Experimental Reading Readiness Tests* (ERR) were administered by the classroom teachers during the second and third weeks of September, 1975. No more than two subtests were administered in any one sitting. The *Canadian Readiness Tests* (CRT), by Evanechko, Ollila, Downing and Braun (1970), were also administered during the same two weeks. During the last two weeks of February, the *Bond-Balow-Hoyt: The New Developmental Reading Tests* (BBH) were administered. All tests were scored by the teachers and then rechecked by the investigators.

TABLE 1  
MEANS AND STANDARD DEVIATIONS ON THE *EXPERIMENTAL READING READINESS, CANADIAN READINESS, AND THE BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS* OF FIRST GRADE CHILDREN

Variable	Boys (n=62)		Girls (n=53)		Both Sexes	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
<u>Experimental Reading Readiness Test</u>						
1. Basic Language Concepts	24.98	2.44	25.58	2.40	25.26	2.42
2. Predicting Outcomes	22.68	3.66	22.26	3.46	22.49	3.54
3. Spatial Relations	22.98	3.53	23.42	3.60	23.18	3.54
4. Auditory Reception	22.45	5.18	22.23	4.02	22.35	4.64
<u>Canadian Readiness:</u>						
5. Technical Language of Literacy	13.21	3.37	13.15	4.12	13.18	3.70
6. Letter Recognition	17.87	3.95	18.21	4.10	18.03	3.99
7. Word Matching	15.26	4.63	17.09	3.75	16.10	4.31
8. Beginning Sounds	16.50	3.34	16.13	3.37	16.33	3.33
9. Semantics	27.76	7.50	30.45	6.77	29.00	7.23
10. Learning Rate	16.65	4.59	17.00	4.56	16.81	4.54
<u>Bond-Balow-Hoyt: The New Developmental Reading Tests:</u>						
11. Word Recognition	19.85	8.45	20.11	8.32	19.97	8.31
12. Comprehending Significant Ideas	15.53	10.84	17.40	10.30	16.39	10.55
13. Comprehending Specific Instructions	12.52	6.10	14.17	6.16	13.29	6.13

Results of the Study

The means and standard deviations are shown in Table 1 for the boys and girls on the *Experimental Reading Readiness Test*, *Canadian Readiness Test*, and the *Bond-Balow-Hoyt: The New Developmental Reading Tests*. On the *Experimental Reading Readiness Test* the means ranged from 25.26 out of 29 items (87.1%) on the Basic Language Concepts subtest to a low of 22.35 out of 30 items (75.0%) on the Auditory Reception subtest.

Sex Differences in Reading  
Readiness and Reading  
Achievement

The girls scored significantly higher than boys on only two subtests. They were the Word Matching subtest, ( $T(113) = 2.29, p < .025$ ) and the Semantic subtest ( $T(113) = 2.01, p < .05$ ) of the *Canadian Readiness Test*. All other differences between sexes on the subtests were not significant,  $p > .15$  in each case.

Reliability of the Independent  
Variables as a Test on the  
Experimental Reading Readiness Tests

The Kuder-Richardson 20 reliability coefficient, which determines the internal consistency of the tests was computed for each of the four subtests on the Experimental battery. Coefficients of the four subtests ranged from .51 to .81. Two of the subtests, Auditory Reception and Spatial Relations Concepts, show a sufficiently high reliability of over .70 (Brunig & Kintz, 1968). The KR-20 reliability coefficient of .88 for the total *Experimental Reading Readiness Test* shows a high reliability when the *Experimental* battery is used as a whole.

TABLE 2  
MEANS, STANDARD DEVIATIONS, RANGES, AND THE  
KUDER-RICHARDSON RELIABILITY COEFFICIENT FOR THE TOTAL  
SAMPLE ON THE *EXPERIMENTAL READING READINESS* SUBTESTS

Variable	Mean	SD	Range	KR-20
X1 Basic Language Concepts	24.69	2.51	17-29	0.505
X2 Predicting Outcomes	21.78	3.94	7-29	0.679
X3 Spatial Relations	23.03	3.58	13-29	0.706
X4 Auditory Reception	22.16	5.10	0-30	0.814
Total Test	91.65	11.40	50-112	0.876

N = 115

The Correlation Analysis

The Pearson Product-Moment Correlation Analysis was conducted to answer hypothesis 3, to determine the extent of the relationships among reading readiness and reading achievement measures. Table 3 shows the

TABLE 3  
INTERCORRELATIONS BETWEEN SUBTESTS ON EXPERIMENTAL  
READING READINESS, CANADIAN READINESS, AND  
BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1.	1.000												
2.	0.372***	1.000											
3.	0.415***	0.466***	1.000										
4.	0.338***	0.349***	0.436***	1.000									
5.	0.281**	0.475***	0.358***	0.185*	1.000								
6.	0.331***	0.307***	0.341***	0.245**	0.536***	1.000							
7.	0.331***	0.432***	0.365***	0.259**	0.372***	0.453***	1.000						
8.	0.432***	0.418***	0.439***	0.362***	0.507***	0.621***	0.402***	1.000					
9.	0.449***	0.397***	0.398***	0.237*	0.415***	0.342***	0.407***	0.330***	1.000				
10.	0.357***	0.363***	0.362***	0.233*	0.462***	0.663***	0.590***	0.660***	0.310***	1.000			
11.	0.322***	0.338***	0.427***	0.264**	0.285**	0.405***	0.384***	0.454***	0.248**	0.574***	1.000		
12.	0.294***	0.220*	0.365***	0.240**	0.229**	0.316***	0.391***	0.358***	0.241**	0.465***	0.799***	1.000	
13.	0.348***	0.285**	0.426***	0.289**	0.262**	0.389***	0.434***	0.444***	0.303***	0.546***	0.834***	0.850***	1.000

\*  $p \leq .05$   
\*\*  $p \leq .01$   
\*\*\*  $p \leq .001$

Readiness and Reading Tests				(Readiness: Numbers 1 to 10; Reading: Numbers 11 to 13)			
1.	Basic Language Concepts	5.	Technical Language of Literacy	9.	Semantics		
2.	Predicting Outcomes	6.	Letter Recognition	10.	Learning Rate		
3.	Spatial Relations Concepts	7.	Word Matching	11.	Word Recognition		
4.	Auditory Reception	8.	Beginning Sounds	12.	Comprehending Significant Ideas		
				13.	Comprehending Specific Instruction		



intercorrelations for the entire sample. A two-tailed test was used to find the *p* significance. All the correlations proved to be significant; however, they varied as to the degree of correlation. *The Bond-Balow-Hoyt: The New Developmental Reading Tests* showed very high correlations between subtest (*r* = .80 to *r* = .85), suggesting that to a large degree they were measuring the same facet of reading achievement. The *Canadian Readiness* subtests showed lower correlations (*r* = .32 to *r* = .66). The correlations between the *Experimental Reading Readiness* subtests were lower (*r* = .35 to *r* = .47) showing a low degree of interrelatedness. One may assume that to a large degree each subtest is measuring a different facet of reading readiness on the *Experimental Reading Readiness Test*.

Multiple Regression Analysis  
to Find the Best Predictors  
of Reading Achievement

Hypothesis 4 was tested by developing and comparing multiple linear regression models to determine which readiness measures were the best predictors of reading achievement. The stepwise regression analysis was employed and the predictor variables entered the regression equation in order of their greatest contribution to the increase in *R*<sup>2</sup>. Tested for significance was each entering variable.

Results showed that the Learning Rate subtest on the *Canadian*

TABLE 4  
MULTIPLE REGRESSION—EXPERIMENTAL READING READINESS,  
CANADIAN READINESS, AND BOND-BALOW-HOYT:  
THE NEW DEVELOPMENTAL READING TESTS (SUBTEST 1)

		Subtest #1		
Step No.	Variable	Word Recognition % of Variance <i>R</i> <sup>2</sup>	Probability Level	Standard Error
1.	Learning Rate	33.56	0.000*	6.839
2.	Spatial Relations	39.12	0.002*	6.576
3.	Predicting Outcomes	39.37	0.494	6.591
4.	Technical Language of Literacy	39.67	0.463	6.605
5.	Basic Language Concepts	39.85	0.570	6.625
6.	Auditory Reception	39.93	0.707	6.652
7.	Beginning Sounds	39.98	0.759	6.680
8.	Semantics	40.01	0.829	6.710
9.	Word Matching	40.01	0.973	6.742
10.	Letter Recognition	40.01	0.993	6.774

N = 115

TABLE 5  
MULTIPLE REGRESSION—*EXPERIMENTAL READING READINESS, CANADIAN READINESS, AND BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS (SUBTEST 2)*

Subtest #2				
Step No.	Variable	Comprehending Significant Ideas % of Variance R <sup>2</sup>	Probability Level	Standard Error
1.	Learning Rate	21.99	0.000*	0.396
2.	Spatial Relations	26.48	0.010*	9.162
3.	Word Matching	27.55	0.204	9.137
4.	Basic Language Concepts	27.91	0.458	9.155
5.	Predicting Outcomes	28.36	0.412	9.169
6.	Auditory Reception	28.59	0.554	9.196
7.	Letter Recognition	28.74	0.634	9.229
8.	Technical Language of Literacy	28.77	0.844	9.270
9.	Beginning Sounds	28.78	0.868	9.314
10.	Semantics	28.79	0.904	9.358

N = 115

*Readiness Test* was the best predictor of reading achievement on all three subtests of the *Bond-Balow-Hoyt: The New Developmental Reading Tests*. The *Spatial Relations* subtest on the *Experimental Reading Readiness Test* was also a significant predictor of reading achievement ( $p < .01$  to  $p < .001$ ). A second significant predictor of the *Bond-Balow-Hoyt*: subtest *Comprehending Specific Ideas* within the *Experimental Battery* was the *Basic Language Concepts* subtest. However, this occurred only in the absence of the *Canadian Readiness Test* from the predictor set ( $p < .025$ ).

Discussion

The investigator's *Experimental Reading Readiness Test* had a high coefficient of reliability as a total test, indicating that it could be useful in testing for reading readiness, particularly if included with other testing.

Very few significant differences were found in reading readiness and reading achievement between boys and girls. There was no significant differences in reading readiness on the investigators' *Experimental Reading Readiness Test*. The girls, however, showed a significant difference in reading readiness on two of the subtests on the *Canadian Readiness Test*, *Word Matching*, and *Semantics*. There were no significant differences in reading achievement between boys and girls.

The subtest *Spatial Relations Concepts* on the *Experimental Reading*

TABLE 6  
MULTIPLE REGRESSION—*EXPERIMENTAL READING READINESS, CANADIAN READINESS, AND BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS (SUBTEST 3)*

Subtest #3				
Step No.	Variable	Comprehending Specific Instructions % of Variance $R^2$	Probability Level	Standard Error
1.	Learning Rate	30.33	0.000*	5.162
2.	Spatial Relations	36.38	0.001*	4.954
3.	Word Matching	37.17	0.240	4.945
4.	Basic Language Concepts	37.80	0.295	4.943
5.	Technical Language of Literacy	38.21	0.398	4.949
6.	Auditory Reception	38.52	0.461	4.960
7.	Semantics	38.66	0.620	4.977
8.	Beginning Sounds	38.80	0.624	4.995
9.	Predicting Outcomes	38.91	0.668	5.014
10.	Letter Recognition	38.93	0.848	5.037

N = 115

*Readiness Test* proved to be the best predictor of reading achievement on all three subtests of the *Bond-Balow-Hoyt*. Basic Language Concepts was a significant predictor on the one subtest of the BBH-Comprehending Specific Instructions again, however, in the absence of the *Canadian Readiness Test*. It appears that both these subtests would be valuable contributors on measurements to predict reading. On the *Canadian Readiness Test*, Learning Rate proved to be a very strong predictor of reading achievement on all three subtests of the BBH. On this subtest the teacher teaches ten sight words. The child is then required to select from a row of three words one of the specific words named by the teacher. This subtest is designed to measure the child's ability to learn sight words, a very major part of learning to read.

From the findings of this study, Learning Rate and Spatial Relations Concepts appear to be important contributors to the prediction of reading readiness. Other variables, however, should not be discarded too soon. If the difficulty level of the *Experimental Reading Readiness Tests* had been increased and reliability improved, it is possible that the other subtests may have made a significant contribution as predictors of reading achievement.

Girls did not prove to be superior to boys in reading achievement, as had been found by Dykstra and Tinney (1969) where girls were found to be



superior in reading achievement. This may not have remained true if the reading tests had been administered in June instead of February. Further investigation into differences in reading readiness and reading achievement between boys and girls is required. Instructional needs of boys and girls should also be investigated.

More longitudinal research in reading achievement should be undertaken to clarify the abilities necessary for beginning reading and improve instruments to measure these abilities. Another area for research would be analyzing what readiness skills are important in a beginning reading program using an initial synthetic approach to word recognition as compared with a programme based on a whole word approach. Further research would also be useful in developing readiness tests which diagnose the best modality for the child's initial reading instruction.

#### References

- Adelman, H. S., & Feshbach, S. Predicting reading failure: Beyond the readiness model. *Exceptional Children*, 1971, 37 (5), 349-354.
- Barrett, T. C. Visual discrimination tasks as predictors of first-grade reading achievement. *The Reading Teacher*, 1965, 18, 276-282.
- Bond, G., Balow, C., & Hoyt, C. *Bond-Balow-Hoyt: The New Developmental Reading Tests*. Chicago: Lyons & Carnahan, Inc., 1965.
- Brunig, J. L., & Kintz, B. L. *Computational handbook of statistics*. Glenville, Illinois: Scott Foresman, 1968.
- de Hirsch, K. Tests designed to discover potential reading difficulties at the six-year-old level. *American Journal of Orthopsychiatry*, 1957, 27, 566-576.
- de Hirsch, K., Jansky, J., & Langford, W. *Predicting reading failure*. New York: Harper & Row, 1966.
- Dykstra, R., & Tinney, R. Sex differences in reading readiness—First grade achievement and second grade achievement. In A. J. Figurel (Ed.), *Reading and realism*. Newark, Delaware: International Reading Association, 1969, 623-628.
- Evanechko, P., Ollila, L., Downing, J., & Braun, C. *Canadian Readiness Test*. Unpublished, University of Victoria, 1970.
- Gates, A. I., & Bond, G. Reading readiness: A study of factors determining success and failure in beginning reading. *Teachers College Record*, 1936, 37, 679-685.
- Goodacre, E. *Children learning to read*. London: Routledge & Kegan Paul, 1971.
- Johnson, R. E. The validity of the Clymer-Barrett Pre-reading Battery. *The Reading Teacher*, 1969, 22, 69-75.
- MacGinitie, W. H. Evaluating readiness for learning to read: A critical review and evaluation of research. *Reading Research Quarterly*, 1969, 4, 396-410.
- Ohnmacht, D. D. The effects of letter-knowledge success—First grade. Paper presented at the annual A.E.R.A. Convention, Los Angeles, 1969.
- Popham, W. J., & Sirotnik, K. A. *Educational statistics: Use and interpretation*. New York: Harper & Row, 1973.
- Samuels, S. J. Letter-name versus letter-sound knowledge in learning to read. *The Reading Teacher*, 1971, 24, 604-608; 662.
- Spache, G. D., & Spache, E. B. *Reading in the elementary school* (2nd ed.). Boston: Allyn and Bacon, Inc., 1971.
- Standish, E. J. Readiness to read. *Educational Research*, 1959, 12, 29-38.
- Weintraub, S. Readiness measures for predicting reading achievement. *The Reading Teacher*, 1967, 20, 551-558.

BETTY HORODEZKY

*The University of British Columbia*

## Comparative Difficulty of Beginning Reading Vocabulary: Set II

*A second set of 50 most frequently occurring words found in early primary reading materials was taught to 108 kindergarten pupils through the use of one of three instructional methods: the phonic, the kinesthetic, or the meaningful context. The rank order of difficulty for the 50 words was determined through ease of learning indices in the form of percentages based on the total number of correct responses recorded for each word. Data derived through use of the Spearman rank order correlation coefficient indicated that a nonsignificant relationship existed between rank order of word difficulty indices and rank order of frequency of occurrence for the words when taught by the three methods.*

*This study also included rank order indices for 13 additional words replicated from the first set of most frequently occurring words previously studied by Mangieri in an earlier investigation of this series. The rank order of difficulty indices reported in the present study for the same words was found to be not significantly different.*

Numerous reading authorities agree that learning to read requires the child to develop an adequate sight vocabulary of meaningful words prior to instruction in word analysis (Chall, 1967). Despite those who express divergent opinions, the advantages of acquiring an early sight vocabulary are apparent for a number of reasons (Smith, 1963).

First, sight vocabulary development is necessary because it serves to facilitate the learning of difficult words which do not lend themselves to regular phonic analysis or structural word attack methods; second, sight words permit the child to focus on meaning rather than on "decoding"; third, there is greater immediate intrinsic value in learning to read whole words instantaneously than in learning to read letters in preparation for

---

Betty Horodezky received her Ph.D. in Reading and Language Arts from the University of Pittsburgh, where she served as Research and Teaching Associate. She has taught on the faculties of universities in Alberta, Colorado, Pennsylvania, and Nova Scotia. Dr. Horodezky has contributed to various IRA publications. She is currently a member of the Reading Department at the University of British Columbia.



reading (Anderson, 1968); and fourth, some sight words serve to provide basic understandings in phonetic and structural analysis which later enable the child to attach sound and meaning concepts to many unfamiliar words encountered in reading (Gray, 1960).

Regardless of whether a sight word approach or a word analysis approach is employed, initially, a decision concerning appropriate word selection for beginning reading must be made. Because of this need, numerous primary grade vocabulary lists have been developed over the past five decades. Although the authors of these works have rendered a significant service to reading, several major issues appear to have been overlooked in their quest to develop these lists. First, few significant studies were developed to investigate the relative learning difficulty of individual words contained within a given word list. Second, until recently only a small number of studies had been conducted to determine which sight words should be introduced first.

The present study, the second in a series of word studies, was thus initiated for the major purpose of more closely examining these problems (Sartain, et al., 1968, 1975). This investigation was designed to determine the relative leaning difficulty of a second set of 50 most frequently occurring words found in textbooks and library books for early primary level children, and to reassess the relative learning difficulty of 13 additional words previously studied by Mangieri (1972) and Mangieri and Sartain (1973), in a first set of 63 words taken from the Comprehensive Reading Vocabulary Primary Level (CRV-PL) list (Sartain, 1968-1975). The 13-word replication was done to ascertain consistency of findings. This investigation also included studies of the relationship between word frequency according to the CRV-PL list, the rank order of learning difficulty for the 50 words, and the relative effectiveness of the three methods of vocabulary instruction.

### *Method*

#### *Subjects*

The population of the study consisted of 199 kindergarten children. Since the study was based on the acquisition of unfamiliar sight words, it was necessary that the subjects included in the investigation be limited to those who had not yet learned to read. The Slosson Intelligence Test (SIT) was administered to provide a stratified random sampling whereby subjects were placed into I.Q. categories of high, medium, or low ranges based on their respective SIT scores. Through random selection from within each I.Q. category, nine equivalent groups of 12 children each (4 high, 4 medium, and 4 low I.Q. range subjects) were provided to form a total sample of 108 subjects.

Through random selection, the 63 sight words were divided into three lists (A, B, and C) each containing 21 words. Three instructional methods—the phonic, kinesthetic, and meaningful context—were used to teach these words to the nine equivalent groups of subjects randomly assigned to Sections I, II, and III.

#### *Procedure*

Each section of three groups studied one word list of 21 words on a schedule of 7 words per week. In six weeks, each of the specified groups in a



given section was thus taught 21 words from either word list A, B, or C, utilizing one of three designated methods.

Each sample group of 12 was split into 2 groups of 6 in order to facilitate instruction. Teaching sessions consisted of 15-minute periods on 2 consecutive days in which seven words were taught and retaught by one of the three prescribed methods. Tests of recall were administered immediately following each instructional session and a delayed recall test was given approximately 24 hours after the second instructional session. All testing and teaching instruction was conducted by the writer who had extensive teaching experience at the primary level.

Analysis

The rank order of difficulty for the 63 words was determined through ease of learning indices in the form of percentages. These were based on the total number of correct responses recorded for the three tests of recall for each word. A comprehensive rank order index was conducted by totalling pupils' correct response scores for each word tested following instruction by all methods. The Spearman rank order correlation coefficient was used to determine the relationship between word difficulty and word frequency for the words studied. A comparison was made between the frequency of each word selected from the CRV-PL list, and the total difficulty index.

The effectiveness of the three instructional methods in the acquisition of word lists A, B, and C was assessed in a block treatment analysis of variance design. Post hoc comparisons were made through application of Scheffé's Method of Multiple Contrasts.

In order to determine if there was consistency in the rank order of difficulty indices findings between the first study and the present study for the 13 words replicated, the Spearman rank order correlation coefficient was applied. Table 1 provides data on the comparative rank order of learning difficulty indexes for these words.

TABLE 1  
COMPARATIVE RANK ORDER OF LEARNING DIFFICULTY INDEXES  
FOR THIRTEEN WORDS REPLICATED FROM STUDY I

Present Study	Study I
Rank Order Ease of Learning for Thirteen Words Replicated	Rank Order Ease of Learning for Same Thirteen Words
1. as	1. see
2. look	2. look
3. see	3. are
4. down	4. down
5. at	5. all
6. and	6. she
7. she	7. as
8. not	8. not
9. are	9. at
10. all	10. and
11. how	11. how
12. what	12. then
13. then	13. what

Findings and Conclusions

It was found that a nonsignificant relationship existed between the rank order of word difficulty indices and rank order of frequency of occurrence for the 50 words when taught by using the three methods. These results suggest the frequency of occurrence of words found in primary level reading materials is not related to the learning difficulty of the words. The phonic, kinesthetic, and meaningful context instructional methods were not found to be significantly different in effecting acquisition of word lists A, B, and C. Also, the rank order of learning difficulty for the 50 words was not significantly different for any single method, as illustrated in Table 2. It can be inferred from these findings that the relative learning difficulty of sight words is not significantly affected by teaching method. A moderate to high correlation coefficient was found to exist between the rank order of difficulty indices for the 13 words replicated in the present study with their rank order of difficulty indices reported in the first investigation of this series.

TABLE 2  
RANK ORDER CORRELATION  
BETWEEN FREQUENCY OF OCCURRENCE  
AND WORD DIFFICULTY INDEXES FOR WORD SET II (50 words)

Variables	2	3	4	5
1. Phonic Rank Order of Difficulty Index	.77*	.67*	.89*	-.15
2. Kinesthetic Rank Order of Difficulty Index		.76*	.94*	-.05
3. Meaningful Context Rank Order of Difficulty Index			.88*	.05
4. Comprehensive Rank Order of Difficulty Index				-.06
5. Frequency of Occurrence Rank Order Index				

\* Significant at the .05 level.

Implications for Teaching

It might be inferred from the results of this study that words most commonly appearing in early primary reading materials are not of equal difficulty; some words appear to be more easily learned than others, regardless of their frequency of occurrence or of the method by which they are taught, as shown in Table 3.

In view of these findings, it is suggested that authors of primary reading materials and classroom teachers engaged in beginning reading instruction place greater emphasis on the relative learning difficulty of words and on the order in which these words are presented. Selection of appropriate words according to their ease of learning indices might well provide for the initial success so vital to the beginning reader, for "Among educators, parents, or psychologists, there would be few dissenters to the propositions that 'getting the right start' in learning to read is of greatest importance" (Heilman, 1972, p. 27).

TABLE 3  
COMPREHENSIVE RANK ORDER OF DIFFICULTY FOR FIFTY WORDS  
(SET II) INCLUDING THE THREE METHODS TAUGHT

Comprehensive Rank Order of Words Learned	Ease of Learning Percentage	Rank Order of Frequency of Occurrence--CRV-PL List	Comprehensive Rank Order of Words Learned	Ease of Learning Percentage	Rank Order of Frequency of Occurrence--CRV-PL List
1. too	70.30	10	26. where	22.45	37
2. Mr.	70.00	26	27. time	19.00	42
3. cat	54.46	43.5	28. going	18.27	34
4. if	53.06	25	29. tens	17.82	15
5. house	52.48	36	30. want	15.84	12
6. back	49.00	14	31. long	15.00	50
7. her	48.00	5	32. can't	14.42	46
8. know	47.57	8	33.5 something	14.29	23
9. help	41.84	27	33.5 from	14.29	1
10. man	39.42	17	35.5 were	14.00	2
11. by	39.00	39	35.5 some	14.00	28
12. red	38.84	41	37. find	13.27	13
13. or	37.62	38	38. has	13.00	35
14.5 things	37.00	48	39. day	12.00	29
14.5 don't	37.00	49	40. got	11.88	47
16. two	33.66	31	41. words	11.00	16
17. ran	33.01	18.5	42. tell	10.89	24
18. name	33.00	4	43. into	9.71	22
19.5 make	31.68	6	44. just	8.74	7
19.5 good	31.68	11	45. right	7.92	32
21. saw	31.00	33	46. each	6.80	21
22. play	29.70	43.5	47.5 them	4.95	9
23. write	26.00	40	47.5 put	4.95	20
24. came	24.49	30	49. ones	3.85	45
25. many	23.76	3	50. could	0.00	18.5

References

Anderson, V. D. *Reading and young children*. New York: The Macmillan Company, 1968.

Chall, J. *Learning to read: The great debate*. New York: McGraw-Hill, 1967, pp. 13-14.

Gray, W. S. *On their own in reading*. Glenview, Illinois: Scott Foresman and Company, 1960.

Heilman, A. W. *Principles and practices of teaching reading*. Columbus, Ohio: Charles E. Merrill Publishing Company, 1972.

Mangieri, J. N. *Difficulty of beginning sight vocabulary taught by three methods*. Unpublished Ph.D. dissertation. School of Education, University of Pittsburgh, 1972.

Mangieri, J. N., & Sartain, H. W. Difficulty of beginning sight vocabulary taught by three methods. *Child Study Journal*, 1973, 3(3), 107-114.

Sartain, H. W., and others. Unpublished series of vocabulary studies, University of Pittsburgh, 1968, 1972, 1973, 1975.

Sartain, H. W. Unpublished Comprehensive Reading Vocabulary Primary Level (CRV-PL) List. University of Pittsburgh, Falk Laboratory School, 1968-1975.

Smith, N. B. *Reading instruction for today's children*. Englewood Cliffs: Prentice-Hall, Inc., 1963.



JOHN McARTHUR

Monash University

## Teacher Socialization: The First Five Years\*

*The study reported here relates to later stages of a longitudinal project which attempts to replicate in an Australian context the findings observed in North America and England regarding teacher socialization during the early career phase. The study involves a large sample monitored over a 5- to 6-year period and uses as its primary measure of socialization the Pupil Control Ideology Form, which is now well established in educational research. The basic finding is that, after an initial period of trauma, teacher socialization enters a plateau phase wherein those teachers who remain in their chosen occupation commence a process of value internalization.*

How influential are the first five years of teaching? The study reported here attempts to provide some answers to this question by focussing upon one phase of a longitudinal study of beginning secondary teachers in Victoria, Australia. Earlier phases of teacher socialization have been described in terms of (a) "anticipatory socialization" during the teacher training phase, and (b) "reality shock", the phenomenon which occurs when idealism gives way to reality during the initial period of full-time experience. Phases (a) and (b) have been examined in a number of professions—medicine (Becker & Geer, 1958; Hughes, 1956), law (Lortie, 1959), and nursing (Corwin, 1961) are examples. In an earlier article describing phase (b), McArthur (1978) reported observations of "reality shock" with a group of beginning secondary teachers in terms of a movement in their self-reported "custodialism" towards pupils. A sample of 534 beginning teachers responded to statements on pupil control contained in the Pupil Control Ideology Form (PCI Form, Willower, Eidell & Hoy, 1967). By comparison with this group, who had commenced their initial teaching experience in 1973, a smaller group of 30, who had undertaken teacher training at the

---

John T. McArthur, B.Comm., B.Ed., Ph.D., is a Senior Lecturer in Educational Administration, Monash University, Australia. His teaching and research interests are in fields of teacher socialization, pupil control, organizational climate, computer techniques in educational research, multivariate statistical analysis. Dr. McArthur has been engaged in teaching, research, and publication at the university level since 1969 in school curriculum and administration.

\* The study is supported by financial assistance from the Educational Research and Development Committee, Canberra, Australia.

TABLE 1  
COMPARISON OF PCI SCORES BEFORE AND AFTER INITIAL TEACHING  
EXPERIENCE

Groups	<i>N</i>	Pretest PCI		Posttest PCI		<i>t</i>
		Mean	Standard Error	Mean	Standard Error	
Hoy (1968)						
Teaching	58 <sup>a</sup>	51.97	1.05	54.91	1.03	2.87*
Nonteaching	39	47.31	1.55	47.05	1.28	0.27
McArthur (1973)						
Teaching	534	44.51	0.41	50.88	0.44	17.63***
Nonteaching	30	45.57	1.49	45.20	1.69	0.40

<sup>a</sup> To enable more useful comparison, only the secondary teachers in Hoy's study (i.e., half his sample) are included in this table.

Probability levels for *t* values

\*\*\* Significant at .001  
\*\* Significant at .01  
\* Significant at .05

same time as the 534 but had not taught during 1973, displayed no change in their attitudes to pupil control. The findings of this section of the study were supportive of those reported earlier by Hoy (1969). Comparative results of Hoy (1969) and McArthur (1978) are included as Table 1.

Similar observations of the “reality shock” phenomenon in beginning teachers have been reported by Day (1959), using the Minnesota Teacher Attitudes Inventory, and by Morrison and McIntyre (1967) using the Manchester N, R and T scales devised by Oliver and Butcher (1962).

The study reported here is grounded in occupational socialization theory and the associated literature on reference groups (Hartley, 1960; Hyman & Singer, 1968). The suggestion that beginning teachers feel the need to comply with subcultural values of the teaching profession in order to achieve reference group membership is a common one. As Waller (1932) says:

The significant people for a school teacher are other teachers . . . . A landmark in one's assimilation to the profession is that moment when one decides that only teachers are important. (p. 389)

It is also frequently argued that concern with discipline and control of pupils is a dominant subcultural value of teaching. (Waller, 1932; Silberman, 1970). According to Silberman, “The most important characteristic schools share in common is a preoccupation with order and control” (1970, p. 122).

*The Notion of Pupil Control Ideology*

The instrument used in this study as a proxy measure of teachers' self-reported attitudes to pupil control is the PCI Form. Details of the construction and validation of this instrument are contained in the original

monograph. The PCI Form consists of 20 items requiring Likert-type responses, ranging from Strongly Agree to Strongly Disagree. Respondents' scores are a summation of responses to the 20 items, ranging along a continuum from a possible low of 20 (not custodial) to a possible high of 100 (very custodial). Typical items are: "It is important for pupils to learn to obey rules rather than that they make their own decisions"; "Pupils often misbehave in order to make the teacher look bad."

PCI scores of teachers in a wide variety of school systems have been collected and analysed in relation to a range of biographical and other characteristics or attitude measures (e.g., sex, age, teaching experience, dogmatism, innovativeness). Willower (1977) has recently reviewed a large number of studies which have used the PCI Form. It would seem that there is sufficient evidence to justify acceptance of the notion of pupil control as measured by the PCI Form as an important aspect of the teacher subculture.

### *Prior Observations of the Sample of Teachers in this Study*

In the earlier published report of phase (b) of the longitudinal study upon which the paper is based, some observations were made concerning the group of beginning teachers used to provide base-line data for the study. These observations are presented again here in summarized form for comparative purposes.

During 1972 a sample of 800 beginning teacher trainees, representing roughly 50% of the anticipated intake of new secondary teachers to the government school system<sup>1</sup> for the 1973 teaching year, were surveyed. The survey was conducted late in the year, after all practice teaching had been completed. Of the original group of 800, 564 (70.5%) completed a follow-up questionnaire in 1973, about 8 months after the completion of their teacher training courses. The PCI Form was used as the major part of the survey instrument on both occasions. Of the 564 subjects who completed the post-training questionnaire in 1973, 534 had been teaching during the intervening period and 30 had not taught at the secondary school level for various reasons (further study, pregnancy, illness, tertiary teaching appointment).

As reported above, a major finding of the earlier publication on the first stage of this study was the difference between those who had taught and those who had not. This finding supported Hoy's work and suggested a possible relationship between treatment (having been exposed to a teaching situation between pretest and posttest observations) and effect (increase in their custodialism as measured by PCI scores from pretest to posttest).

Differences between subgroups of the samples at both pretest and post-test phases were examined in cross-sectional analyses: e.g., based upon sex (males tended to be more custodial than females on both occasions); teaching subjects (teachers of physical sciences and practical subjects, such as physical education and home economics, were much more custodial than teachers of humanities subjects who, in turn, were more custodial than those prepared for counselling and guidance in schools); and commitment to teaching as a career (committed teachers tended to be more custodial than those uncommitted to a career in teaching). There were also marked differences between the group means of trainee teachers based upon the eight secondary teacher training courses extant in 1972. In most cases, at



the posttest stage these between-group differences persisted in the same rank order, precluding any real conclusions or speculations regarding differential patterns of socialization based upon observable characteristics of a "fixed" type; i.e., sex, teaching subjects, training institution. The exception was commitment, which, unlike the "fixed" characteristics, had a built-in flexibility, or potential for change. Very pronounced changes in PCI scores (between 24 and 33% compared with an overall gain of 13%) were recorded in the case of two of the nine possible situations when the three categories ("strongly committed," "weakly committed," "uncommitted") at both stages, were cross-tabulated. Paradoxically, not only those who moved from a position of weak commitment before teaching to a strong position after initial teaching experience, but also those whose strong pretest commitment had been changed to weak commitment after teaching, recorded strong PCI gains between the two observations. Dr. Itai Zak, Tel Aviv University, who is conducting a large-scale longitudinal study of beginning teachers in Israel (Zak, 1979), involving use of the PCI Form, has suggested a psychological explanation for this apparent paradox. In a recent personal communication he proposed that less stable teachers (i.e. those of high psychological uncertainty about the best course of action to choose to satisfy psychological needs) tend to gravitate towards the custodial end of the PCI Form continuum.

#### *The Extended Longitudinal Phase of the Study*

The present longitudinal study covers a longer period than most relating to the early years of teaching. Hoy (1969) published details of a follow-up of his sample of beginning teachers in their second year of teaching. It is not known whether he has kept track of this sample as they entered the mid-career phase. Wiseman and Start (1965) reported a follow-up study of teachers in England after their first five years of teaching. Walker (1967) conducted a five-year study of teacher commitment using a group of Australian beginning teachers. Charters (1970) touched upon aspects of this career phase in his study of teacher survival. Mason (1961) conducted a major study of beginning teachers in the U.S.A. but his research was mainly confined to the first year of teaching experience. Ortiz (1978) reviewed the topic of mid-career socialization of educational administrators but provided few details of the "make or break" period of the first five years. Becker's (1957) study of the career patterns of the Chicago public school teacher is one of the few detailed analytical studies of this vital career phase.

The present study appears to be one of the few longitudinal projects being carried out that involves a sizeable cohort of beginning teachers who have been monitored from their preservice years through to a stage of relative occupational maturity/security within their chosen profession in a large educational system. The data that are reported below were obtained in 1978 from 456 of the original sample of 800 trainee teachers surveyed in 1972. These 456 respondents can be categorised as follows:

	(a) Teaching in 1978	(b) Not teaching in 1978
1. Those who completed the survey on all three occasions, 1972, 1973 and 1978	278	80
2. Those who completed the survey in 1972, could not be contacted or failed to respond to the 1973 survey, but completed the 1978 survey (this includes some teachers who resigned during the period but have since rejoined the government teaching service)	76	22
Subtotal of 1978 survey respondents	354	102
3. Those who completed the 1972 and 1973 surveys but could not be contacted or failed to respond to the 1978 survey	194	12
4. Those who completed the 1972 survey but have not participated in either the 1973 or 1978 surveys (This includes 1972 participants known to have died or taken up permanent residence outside Australia. Many of these 1972 participants have never been employed in the Victorian government school system according to official records)	20	118
Subtotal of nonrespondents to 1978 survey	214	130
Total sample of respondents to 1972 survey	568	232

For the purposes of this report, categories 1(a) and 2(a) are the most useful in terms of an “experimental” group. The matching categories of 1(b) and 2(b) (comprising those respondents who were not teaching in 1978) provide a basis for comparison, possibly a “control” group.

Certain features of the “experimental” and “control” groups seemed worthy of cross-sectional analysis. These are discussed briefly here prior to consideration of the major findings of the study.

One interesting, somewhat surprising finding of the study concerns the career patterns of the teaching respondents during the period 1973 to 1978. Respondents were asked for details of place and nature of their employment for each year from 1973 to 1978 inclusive. Of those 354 who were still teaching in 1978, a remarkably high proportion (31.9%, i.e., 113 teachers) had maintained unbroken tenure at their school of initial appointment in 1973. 26 teachers resigned during the period 1973 to 1978 but had rejoined the government teaching service and were once again employed in full-time teaching positions in 1978.

The sex ratio of teachers in the sample of 354 in the experimental group reflected a predictable shift over the 5-year period. In the original sample of 800 teacher trainees, there were 480 females (60%) and 320 males (40%). In the experimental group of 354 there were 171 females (48.3%) and 183 males (51.7%). The career orientation of female beginning teachers has been commented upon by both Mason (1961) and Walker (1967), that is, a significant number indicated that they regarded teaching as a contingent career, one which they could move into and out of during their lifetime.



### Analysis of Results

Data were analysed using the techniques of one-way analysis of variance, for cross-sectional comparisons, and *t* tests for correlated samples, for longitudinal comparisons of group means.

#### Cross-sectional analysis

From a simple comparison of the PCI scores shown in Table 2, one would conclude that there was a significant difference between the group means of those who were teaching and those not teaching in 1978. The results support earlier findings contrasting the PCI scores of subjects at the "coal-face" with those not teaching or in administrative situations (Willower, Eidell & Hoy, 1967; Hoy, 1968, 1969; McArthur, 1975, 1978).

TABLE 2  
PCI SCORES OF RESPONDENTS TO 1978 SURVEY

Groups	<i>N</i>	Mean	Standard Error	<i>F</i>
Teaching in 1978	354	52.76	0.53	14.45***
Not teaching in 1978	102	48.53	0.86	
TOTAL	456	51.86		

Probability levels for *F* values

\*\*\* significant at .001

\*\* significant at .01

\* significant at .05

Cross-sectional comparisons were made using three characteristics as bases for grouping: sex, teaching subjects, and expressed preservice commitment to teaching. These are shown in Table 3 along with comparative figures from the 1972 sample, both the total group of 800 and the specific group of 354 respondents who were teaching in 1978.

Earlier observations that males tend to display more custodialism in their PCI scores than females (Willower, Eidell & Hoy, 1967; McArthur, 1978) were confirmed by the findings shown in Table 3. The increased proportion of males in the 1978 survey was commented upon above. The generally accepted phenomenon of secondary teaching as a young, increasingly "female" profession, at least in the Australian context, is not denied. It is suggested however that global figures often conceal some interesting underlying patterns.<sup>2</sup>

Differences in mean PCI scores when groups were based upon teaching subjects were similar to those observed during the "reality shock" phase (McArthur, 1978). Teachers of sciences and practical subjects were much more custodial than those teaching humanities. Those few trained in guidance and counselling but actually teaching in classrooms in 1978 (*n* = 4) were still very "humanistic" compared with all other teachers. Most of the colleagues of teachers in this small group were now either in consultative positions with the counselling and guidance service centres of the educational authority or had left the teaching service for other employment.



TABLE 3  
PCI SCORES OF TEACHING RESPONDENTS TO 1978 SURVEY, CLASSIFIED  
ON BASIS OF SEX, TEACHING SUBJECTS, AND  
PRESERVICE COMMITMENT TO TEACHING

Groups	<i>N</i>	Mean	Standard Error	<i>F</i>
Male	183	53.74	0.75	3.57**
Female	171	51.71	0.76	
Humanities subjects	181	50.97	0.75	9.42***
Sciences and Practical subjects	169	54.94	0.74	
Guidance subjects	4	42.00	4.02	
Strongly committed	186	52.79	0.74	0.08
Weakly committed	94	52.47	1.11	
Uncommitted	74	53.12	1.10	
1978 teaching respondents	354	52.76		

Probability levels for *F* values

\*\*\* Significant at .001

\*\* Significant at .01

\* Significant at .05

Preservice commitment was not found to be a useful predictor of differences in custodialism in the 1978 sample. Strength of commitment tended to be associated with slightly higher custodialism than weak commitment. Lack of commitment was associated with slightly higher custodialism than either of the other two categories. The differences between group means were far too small to make any further comments worthwhile.

*Longitudinal Analysis*

Data analysis in this section was confined to those 358 respondents to the 1978 survey for whom PCI scores had been obtained on each of three occasions (1972, 1973, 1978). While 278 of these were teaching in 1978, 80 had ceased teaching for various reasons: leave for travel overseas; further study; homemaking; pregnancy; secondment as specialist consultants; other occupations.

Table 4 shows comparative scores for the three surveys; the responses of those teaching in 1978 are separated from those not teaching at that time.

The plateau “effect” observed supports the basic notions of theory in occupational socialization. The “reality shock” phase is accompanied by dramatic shifts as idealism is replaced by realism. Thereafter, a flattening of the curve occurs as stability and security in the occupation consolidate attitudes.

Hoy (1969) observed a difference between those teachers who moved to a new school in their second year of teaching. In contrast to those who remained at the same school and whose PCI scores in Year 2 were

TABLE 4  
COMPARISON OF PCI SCORES OF 1978 RESPONDENTS WHO ALSO  
RESPONDED TO 1972 AND 1973 SURVEYS

Groups	N	1972 PCI		1973 PCI		1978 PCI	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Teaching 1978	278	45.45	0.58	52.24	0.61	52.72	0.61
Nonteaching 1978 <sup>a</sup>	80	43.75	0.87	47.89	1.10	48.18	0.98
		F value 2.04*		F value 11.48***		F value 13.02***	
	358	45.07		51.27		51.70	

<sup>a</sup> A number of those not teaching in 1978 were teaching in 1973. This accounts for the apparent contradiction between these figures and Table 1. In fact only 18 of the 30 "nonteaching" respondents represented in Table 1 are included in the 80 "nonteaching 1978" in Table 5.

Probability levels for F values

- \*\*\* Significant at .001
- \*\* Significant at .01
- \* Significant at .05

TABLE 5  
COMPARISON OF PCI SCORES IN 1972, 1973, AND 1978 SURVEYS OF 1978  
TEACHING RESPONDENTS, CLASSIFIED ON THE BASIS OF  
STABILITY OF TENURE

Groups	N	1972 PCI		1973 PCI		1978 PCI		Correlated t tests		
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	1972:73	1972:78	1973:78
Still at same school as in 1973	91	45.45	1.03	52.71	0.94	52.59	0.98	7.58***	6.31***	0.09
Taught in 2 or more schools	143	46.04	0.80	52.54	0.87	53.31	0.88	9.67***	8.07***	1.06
Resigned between 1973-77 but rejoined the service	44	43.52	1.50	50.30	1.75	51.05	1.64	3.98***	4.01***	0.75
		F value: 1.04		F value: 0.80		F value: 0.26				
Total	278	45.45		52.24		52.72		13.49***	11.57***	0.82

Probability levels for F and t values: \*\*\* significant at .001  
\*\* significant at .01  
\* significant at .05

practically identical with those in Year 1, the mobile teachers seemed to become more custodial. Hoy postulated a “resocialization” phenomenon as the underlying cause. As noted above a sizeable proportion of the experimental group in the 1978 survey were at the same school from 1973 to 1978 thus it was considered appropriate to subdivide the group on the criterion of stability of tenure. The 278 teachers in the experimental group

fell into 3 distinct categories for the purpose of this section of the analysis: those who had taught at 2 or more schools during the period; and those who had resigned but later rejoined the teaching service during the period. Comparative figures are shown in Table 5.

The trends shown here are interesting although not statistically significant. A surprisingly high proportion of teachers have remained at their schools of initial teaching appointment (91, or 32.7% of the 278 still teaching in 1978 and who responded to all three surveys) and the mean PCI score of this group is lower than that of the 143 who have taught at two or more schools in the same period. This high rate of stability of tenure suggests an interesting side issue with respect to career patterns. There may be some relationship between increasing teacher mobility and lack of promotion opportunities in the situation of teacher over-supply common in most Western industrialized countries. The planned follow-up interviews with approximately one-third of the 1978 survey respondents may throw further light upon this issue.

A further set of longitudinal analyses was undertaken using the three variables isolated for cross-sectional purposes in Table 3, namely, sex, teacher training subjects and expressed preservice commitment to a career in teaching. Subgroup PCI means for 1972, 1973 and 1978 are shown in Table 6. As reported in earlier discussions comparing the 1972 and 1973 data (McArthur 1975, 1978), the response patterns over time of subgroups classified on these three bases are not sufficiently dissimilar to allow claims for possible underlying causal factors.

TABLE 6  
COMPARISON OF PCI SCORES IN 1972, 1973 AND 1978 SURVEYS OF 1978  
TEACHING RESPONDENTS, CLASSIFIED ON BASIS OF SEX, TEACHER  
TRAINING SUBJECTS, AND EXPRESSED PRESERVICE COMMITMENT TO  
TEACHING AS A CAREER

Groups	N	1972 PCI		1973 PCI		1978 PCI		Correlated t tests		
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	1972:73	1972:78	1973:78
Male	143	46.41	0.88	53.07	0.85	53.38	0.83	8.70***	7.40***	0.49
Female	135	44.42	0.75	51.36	0.89	52.02	0.90	10.68***	9.17***	0.85
		F value: 2.92*		F value: 1.24		F value: 1.24				
Humanities subjects	141	43.48	0.81	51.21	0.84	50.78	0.86	10.82***	8.50***	0.54
Sciences and practical subjects	133	47.87	0.81	53.58	0.89	55.09	0.84	7.96***	7.61***	1.72
Guidance subjects	4	34.25	3.23	44.00	3.72	42.00	4.02	3.51*	4.31*	0.85
		F value=10.31***		F value=3.23**		F value=8.30***				
Strongly committed	147	45.98	0.81	52.54	0.82	52.25	0.85	9.55***	8.06***	0.40
Weakly committed	75	45.43	1.02	52.83	1.16	53.39	1.21	9.04***	6.50***	0.48
Uncommitted	56	44.07	1.42	50.56	1.47	53.04	1.32	4.69***	5.16***	1.60
		F value: 0.88		F value: 0.58		F value: 0.25				
Total	278	45.45		52.24		52.72		13.49***	11.57***	0.82

Probability levels for F and t values: \*\*\* significant at .001  
\*\* significant at .01  
\* significant at .05



Further analysis is being undertaken to establish whether differential patterns may be more discernible using multivariate rather than univariate procedures. Some interesting differences persist over time in the data when the 1972 training course is used as the basis for grouping respondents. Furthermore, there are some marked differences observable when these institution-based groups are further subdivided on the basis of sex.

Another area of continuing interest is the relationship between changes in expressed commitment to teaching as a career over a period of the three surveys and the extent and direction of changes in custodialism as measured by self-reported PCI scores in 1972, 1973 and 1978. This issue will be examined in a further report on the data, incorporating follow-up interviews.

### Conclusion

The analyses reported in this article indicate that the most marked adjustment in the socialization of beginning teachers takes place during the period of initial teaching experience. This period of "reality shock" is followed by a consolidation or plateau phase during which those who choose to remain in the teaching profession adjust and "internalize" values of the teaching subculture. There is slight, perhaps insufficient, evidence in the data to suggest that the plateau is not so level for those who experience some instability during this phase of "midcareer socialization." Such instability may result from movement to a new school involving "resocialization," new identification, perhaps even compliance with a new reference group. Or it may result from moving out of teaching for a year or two then back into the profession.

As yet the data analysed are insufficient to enable firmer conclusions regarding these matters. Obviously an underlying bias is present in terms of self-selection of respondents who have chosen to remain in teaching. The high proportion of respondents shown in Table 3 as having expressed strong commitment (196, or 52.5% of the 354 teaching respondents in 1978) when compared to the corresponding number in 1972 (336, or 42% of the total of 800) bears testimony to this underlying bias. Despite this potential for distortion the results that have been analysed appear to support the notion that the post "reality shock" years are much less traumatic than the period of initial teaching experience.

### Notes

1. The government school system accounts for about 80% of the pupils, teachers and schools at the secondary school level (years 7 to 12). In 1973 there were 243,800 pupils, 17,00 teachers, and 354 schools at the secondary level in the Victorian government school system.
2. For example, the median age of secondary teachers in Victoria is 28.9 years. Of the 2,930 secondary teachers under the age of 25, 2,089 (71.3%) are female. The age group 25 to 29 years comprises 5,541 secondary teachers, 3,060 (55.2%) of whom are female. Of the remaining 5,984 secondary teachers in 1978 who were aged 30 years and above, only 2,629 (37.9%) were female. The median age of 28.9 corresponds roughly to the "mid-career" phase, 5 to 6 years after training, that is the basic focus of this paper.

### References

- Becker, H. S. The career of the Chicago public schoolteacher. *American Journal of Sociology*, 1957, 63, 470-477.
- Becker, H. S., & Geer, B. The fate of idealism in medical school. *American Sociological Review*, 1958, 23, 50-56.

- Charters, W. W. Some factors affecting teacher survival in school districts. *American Educational Journal*, 1970, 7, 1-27.
- Corwin, R. The professional employee: A study of conflict in nursing roles. *American Journal of Sociology*, 1961, 66, 604-615.
- Day, H. P. Attitude change of beginning teachers after initial teaching experience. *Journal of Teacher Education*, 1959, 18, 326-328.
- Hartley, R. G. Relationship between perceived values and acceptance of a new reference group. *Journal of Social Psychology*, 1960, 51, 181-190.
- Hoy, W. K. The influence of experience on the beginning teacher. *School Review*, 1968, 76, 312-323.
- Hoy, W. K. Pupil control ideology and teacher socialization: A further examination of the influence of experience on the beginning teacher. *School Review*, 1969, 77, 257-265.
- Hughes, E. C. The making of a physician: A general statement of ideas and problems. *Human Organisations*, 1956, Winter, 21-25.
- Hyman, H. H., & Singer, E. (Eds.) *Readings in reference group theory and research*. New York: The Free Press, 1968.
- Lortie, D. C. Laymen to lawmen: Law schools, careers and professional socialization. *Harvard Educational Review*, 1959, 20, 352-369.
- Lortie, D. C. *Schoolteacher*. Chicago: University of Chicago Press, 1975.
- Mason, W. S. *The beginning teacher: Status and career orientations*. Final Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1961.
- McArthur, J. T. *Teacher socialization: The influence of initial teaching experience on the pupil control ideology of beginning secondary teachers*. Unpublished doctoral dissertation, Monash University, 1975.
- McArthur, J. T. What does teaching do to teachers? *Educational Administration Quarterly*, 1978, 14, 89-103.
- Morrison, A., & McIntyre, D. Changes in opinions about education during the first year of teaching. *British Journal of Social and Clinical Psychology*, 1967, 6, 161-163.
- Ortiz, F. I. Midcareer socialization of educational administrators. *Review of Educational Research*, 1978, 48, 121-132.
- Oliver, R. A. C., & Butcher, H. J. Teacher attitudes to education: The structure of educational attitudes. *British Journal of Social and Clinical Psychology*, 1962, 1, 56-59.
- Silberman, C. E. *Crisis in the classroom*. New York: Random House, 1970.
- Walker, W. Occupational commitment in a sample of Australian teachers five years after training. *Australian Journal of Higher Education*, 1967, 3, 20-32.
- Waller, W. *The sociology of teaching*. New York: Wiley, 1932.
- Willower, D. J. Schools and pupil control. In D. A. Erickson (Ed.), *Educational organization and administration*. Berkeley: McCutchan, 1977.
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology*. Penn State Studies Monograph No. 24. University Park: The Pennsylvania State University, 1967.
- Wiseman, S., & Start, K. B. A follow-up of teachers five years after completing their training. *British Journal of Educational Psychology*, 1965, 35, 342-361.
- Zak, I. (1979) Developmental transitions in the professional growth of teachers. In *Teacher Effectiveness: Proceedings of the 4th Conference of the International Association of Educational Assessment*. New York: Pergamon Press, 1979.

J. M. KIRMAN

and

J. GOLDBERG

*The University of Alberta*

## Student Teacher Telephone Conferencing With Satellite Maps as a Monitoring Device\*

*This exploratory study was conducted to determine the feasibility of telephone conferencing for student teaching supervision. It was found that student teachers performed as well with this mode of supervision as those supervised face to face. However, the student teachers expressed dissatisfaction with the procedure.*

As faculty budgets are cut back, the employment of additional staff for supervising student teachers becomes more difficult. The extended practicum of some jurisdictions also places additional demands on the time of faculty members. The use of telephone conferencing is one way of allowing faculty members to supervise larger numbers of student teachers from a central location, rather than travelling to several locations for supervising a limited number of students.

Present technology allows any number of individual telephones to be joined into a conference. Thus, a faculty advisor may link his student teachers in a conference call at a mutually agreed to time. Technological improvements are presently being developed that may enable a central university switchboard to set up conference calls at a cost comparable to ordinary telephone calls (Edmonton Telephones, Note 1).

In this project three criteria were employed in assessing the effectiveness of telephone supervision: teachers' ratings of student teachers; student

---

Dr. J. M. Kirman is a Professor in the Department of Elementary Education, Social Studies area, and director of Project OMEGA for research in remote sensing education.

Dr. Goldberg is an Associate Professor in the Department of Educational Psychology, and research associate with Project OMEGA.

\* This project was funded by Innovative Projects Fund, Alberta Department of Advanced Education and Manpower, Program Services Division. The cooperation of Dr. Al MacKay, Director of the Center for Research in Teaching, University of Alberta, for external evaluation is gratefully acknowledged.



teachers' ratings of their supervision; and judgment based on examination of the knowledge of select pupils of the student teachers. The latter was based upon specially prepared social studies teaching units on Landsat satellite infrared false color maps.

The following questions were raised:

1. How would student teachers supervised with telephone conferencing compare to another group of student teachers supervised in a face to face manner as evaluated by their cooperating teachers?
2. How would pupils taught about Landsat maps by the above two groups compare with each other?
3. How would the student teachers in each group rate their experiences with the two modes of supervision?
4. What are some variables that influence telephone conferencing as a supervisory tool in student teaching?

### *Previous Research*

There appears to be no literature specifically dealing with telephone conferencing in which a group of student teachers and their faculty advisor are on line at the same time. Some reports of telephone conferencing deal with a telephone receiver hooked to a loudspeaker (Stevenson & Otto, 1975). Another deals with conferencing lecture and discussion groups, but not with student teacher supervision (Madden, 1967). In one case, telephone conferencing was used but only with a single student, the faculty consultant, and the cooperating teacher at one time (Dalrymple, 1971). One study also made use of this "triad" procedure on a single telephone line. Amplifying devices were used by the parties to allow two or more persons to hear or speak with the same receiver (Smith, 1969).

The University of Wisconsin, Télé-Université, Québec, and the Open University, London, England, among others, utilize telephone communications for instruction. Usually, groups of students assemble at central locations for telephone loudspeaker lectures with two-way communication. London also has one to one telephone tutorials and group conferences (Williams, Holloway & Hammond, 1975).

The San Diego Schools utilize group telephone assisted instruction for handicapped children, requiring special installations at the students' homes (Carr, 1964). However, the group telephone project being reported here used existing telephone lines and equipment in the students' homes without any modifications.

### *Procedure*

The study was run during both halves of the school year 1977-1978 with The University of Alberta, Faculty of Education Practicum 201. This is an experience in which participants are expected to teach small groups of pupils. It usually consists of five consecutive weeks where the student teacher is in a school for two days per week. Faculty advisors' roles consists mainly of orienting students and cooperating teachers, following the students' progress, and dealing with problems.

The use of Landsat satellite maps was decided upon to measure the

achievement of the student teachers' pupils. These maps are in infrared false color with a scale of 1:1,000,000. They are produced by Landsat satellites in a near polar orbit, 900 km high, which produces images of almost all the earth's surface every eighteen days.

Since this technology is relatively new, it was expected that neither cooperating teachers, student teachers, nor their pupils had any background regarding satellite maps. This provided a topic in which all student teachers began on an equal footing, and made the student teachers more dependent upon their faculty advisor.

Each student teacher had to teach three children, selected by the cooperating teacher, about Landsat infrared maps for one hour each week. The three children selected were those with the highest reading scores in the class. The children were in grades four, five, and six, and were checked for visual problems. Research had shown that children as young as those in grade three could deal with infrared Landsat maps (Kirman, 1977).

The student teachers were to meet before their teaching round for an orientation session. Telephone and control groups were scheduled to meet at different times, and were assigned to different schools. The subjects were 19 randomly selected student teachers, 10 of whom were randomly selected and allocated to telephone-conference supervision, and the remaining 9 allocated to face-to-face supervision. Table 1 provides a description of the subjects on four variables. The data are not available for one student teacher from the telephone group. However, all other data for this subject are available and are therefore included in all subsequent analyses.

The sex ratio of 2 males and 7 females was identical for both groups. Also, the figures for mean number of years at university (3.6, controls; 3.7, subjects) and mean number of years in the Faculty of Education (2 years in control; 2.3 years, subjects) were virtually identical for both groups. The mean age of the telephone group (25.1) was 3.2 years more than the mean age of the control group. A *t* test on the differences between the age means was clearly nonsignificant ( $t=1.17$ ; d.f.=15) so the two groups can be regarded as comparable in age.

For the controls there were five weekly supervision sessions of at least 45 minutes in each of two schools over the five weeks of the project. The frequency and duration of these supervision sessions exceeded what is normally available to Education Practicum 201 students.

In addition to instruction about Landsat maps, the sessions for the controls dealt with topics pertinent to beginning student teaching. Topics

TABLE 1  
COMPARATIVE DATA FOR STUDENT TEACHER GROUPS<sup>a</sup>  
(*N*= 18)

Group	Mean Years University	Mean Years In Education	Mean Age	Number of	
				Male	Female
Control	3.6	2	21.9	2	7
Telephone	2.7	2.3	25.1	2	7

<sup>a</sup> Data not available on one telephone subject.



were introduced both by the faculty consultant and by the students and handled in group discussion, unilateral presentation, or questions and answers.

The same faculty consultant supervised the telephone subject group in one conference hook-up of 45 to 60 minutes on Thursday evenings for the five consecutive weeks of the project. All participants were at home.

After greeting students and outlining the plan for the evening, the faculty consultant invited each to give a brief account of experiences in school that week. The faculty consultant responded to each account and/or invited others to respond. If necessary, the faculty consultant designated a particular person by name to respond.

Some time was always allocated to Landsat teaching. Also, students selected the previous week reported on topics such as discipline in their class or children with learning problems. The faculty consultant tried to promote group discussion on the presented topic. Time was also devoted to problems. Students with problems which could not be brought up in conference were invited to telephone the faculty consultant privately immediately after. The conference terminated with the faculty consultant providing goals for the student teaching and following conference session.

Note that the control group was in two schools, thus giving each other greater peer support, whereas the telephone group was scattered throughout many schools. The telephone group was usually limited to a maximum of forty-five minutes per week with their faculty consultant. A smaller ratio of student teachers to supervisor favoured the control group since separate 45-minute supervisory sessions were held for the controls in each of the two schools. In addition, the control group was not subject to the stringent time limitations applied to the telephone group. Since many features of the supervision favoured the control group, the present study would appear to be a severe test of telephone supervision.

#### *Evaluation By The Cooperating Teacher— The Cooperating Teacher Evaluation Form*

The student teachers were evaluated by the cooperating teachers who employed the Cooperating Teacher Evaluation Form. This is a standard form requiring qualitative evaluative comments under the headings of attendance, punctuality, attitude, preparation, relationship with pupils, and general comments.

At most, student teachers could be separated into two groups on the basis of the cooperating teachers' ratings—those with consistently positive evaluative comments, and those with generally positive evaluations qualified by a reservation expressed about some teaching characteristics.

The frequency of consistently positive ratings was higher in the telephone group. Nine of ten telephone subjects received a consistently positive rating whereas only three of nine controls received a consistently positive rating. A chi-square test showed this difference to be significant at less than the .05 level ( $\chi^2 = 4.33$ ; d.f. = 1). One can conclude that the telephone group received better global evaluation from their cooperating teachers than did the control group.

A possibility exists that the telephone group's mode of supervision may



TABLE 2  
RATINGS ON COOPERATION TEACHER EVALUATION FORM  
FOR TELEPHONE AND CONTROL SUBJECTS  
( $N=19$ )

Rating	Group	
	Telephone	Control
Positive	9	3
Qualified Positive	1	6

$$\chi^2 = 4.33; \text{ d.f.} = 1; p < .05.$$

have made them more dependent upon the cooperating teachers. This in turn may have fostered a closer relationship which could have had a bearing on the higher evaluations received by the subjects.

The question can also be raised as to whether either mode of supervision resulted in greater student teacher effectiveness than would have been produced with no supervision at all. This cannot be answered by the present data because an unsupervised control group was not included for two reasons: apprehension of this arrangement; and the newness of Landsat material requiring somebody to monitor the student teachers.

Assuming that supervision did make a difference, it can be concluded, at the very least, that telephone supervision would not appear to be poorer than traditional supervision in terms of student teacher effectiveness as assessed by cooperating teachers in this project.

#### *Evaluation Of Pupil Achievement in Landsat Map Reading-- The Pupil Map Questionnaire (P.M.Q.)*

The P.M.Q. consisted of 15 questions based on two Landsat maps of Alberta which the pupils perused in responding to the questions. One part, eliciting a total of 16 responses, required the pupil tested to point to various features, e.g., find a lake. As the child pointed, the examiner noted whether the response was right or wrong. The other part of the P.M.Q. focused on more general knowledge and permitted less objective scoring. In view of the differences between the two parts, they were separated for analysis.

The P.M.Q. was administered individually to the 30 pupils taught by the telephone supervised group and to the 23 pupils taught by the control group of student teachers.

As shown in Table 3 the pupils of the telephone group obtained a higher mean score on the 16-item section (12.9 vs. 11.87) requiring them to point out specific locations on Landsat maps. The difference between the means is statistically significant at the .05 level on a one-tail  $t$  test. On the appropriate two-tail test (since direction of difference was not predicted),  $t$  approaches significance ( $p < .10$  level).

As seen in Table 4, all pupils achieved at virtually identical levels on the general questions (pupils of telephone group, 6.14; pupils of controls, 6.21). High means suggest that both groups may have attained their ceilings on

TABLE 3  
COMPARISON OF PUPILS' ACHIEVEMENT ON  
16-ITEM PUPIL MAP QUESTIONNAIRE  
(*N* = 53 pupils)

	Telephone Ss	Pupils of Control Ss
$\bar{X}$	12.9	11.87
S.D.	1.87	2.11

$t = 1.68$ ; d.f. = 51  
 $p < .10$  (two-tailed test,

TABLE 4  
COMPARISON OF PUPILS' ACHIEVEMENT ON  
7-ITEM PUPIL MAP QUESTIONNAIRE  
(*N* = 51 pupils)<sup>a</sup>

	Telephone Ss	Pupils of Control Ss
$\bar{X}$	6.14	6.21
S.D.	.901	1.099

<sup>a</sup> The responses of two of 53 pupils to the verbal questions were lost due to problems in tape recordings.

$t = .244$ ; d.f. = 49  
N.S. (of two-tailed test)

this kind of test in Landsat knowledge. The  $t$  test on the difference between these means was nonsignificant ( $t = .244$ ).

The data support the conclusion that telephone supervision is clearly not associated with a lesser degree of teaching effectiveness in student teachers.

Since supervision was the main source of Landsat teaching knowledge and techniques for the student teachers, the data offer some support for the conclusion that telephone supervision is at least as adequate as face-to-face supervision in transmitting knowledge about Landsat teaching.

#### *Self-Evaluation of Supervision— The Supervision Evaluation Form (S.E.F.)*

The student teachers' evaluation of their mode of supervision was determined by their responses to the Supervision Evaluation Form (S.E.F.) which was completed anonymously after the final supervisory experience. The items covered aspects of the supervision ranging from its usefulness in providing feedback on their performance to its usefulness in stimulating more thought about student teaching. Content validation is the only method of validity determination for the S.E.F. This is assumed by the form's

construction which consists of systematically delineating functions supervision is intended to serve, and formulating items corresponding to each function.

All eight controls who returned their S.E.F. forms expressed acceptance (ratings of excellent, very good, or good) of face-to-face supervision, whereas only one of nine responding telephone subjects expressed overall acceptance of telephone supervision.

Only two of nine telephone subjects reported telephone supervision to be acceptable, whereas only one of nine would select telephone supervision, if given a choice. Responses indicated that if a choice had to be made between telephone supervision once weekly versus face-to-face supervision once or twice every five weeks, seven of nine subjects would prefer face-to-face supervision.

The results suggest that the telephone group perceived a deficiency in interaction with others during supervision and that the supervision was not sufficiently individualized. Other reasons for dissatisfaction included one expressed frequently about the "impersonal" nature of contact over the phone and the lack of "direct" contact: "It's nice to sit next to someone and ask them how it's going", "would be nice to see people", etc. Complaints expressed no more than once by a single person included poor timing of the call, and the unavailability of the home phone which was needed for business.

Most controls indicated they *would* reject telephone supervision even when the alternative was infrequent face-to-face supervision. Five of the controls spontaneously referred to the disadvantages of telephone supervision although they were not required to contrast their method with telephone supervision! One control noted, "It would be an unpleasant ordeal to sit with an ear to the phone for 45 minutes." These negative references seem remarkable since the controls did not experience telephone supervision, nor have contact with the telephone subjects. This suggests that the controls may have had preconceived negative attitudes toward telephone supervision, and that people generally, including the telephone subjects, may have shared this negative attitude. One of the challenges in using this supervisory mode, therefore, would be to overcome this attitude.

The telephone and control groups did not differ on a majority of the S.E.F. specific items dealing with objectives of supervision such as learning about Landsat, learning from others, obtaining new ideas, being stimulated to think. One can, therefore, venture the conclusion that telephone supervision was functional for imparting important information to student teachers in general, though it succeeded less well in responding to individual needs. Also, a majority of the telephone group listed as an advantage the convenience of telephone supervision.

### *Conclusions*

Student teachers supervised by telephone conferencing were rated as performing better by their cooperating teachers than those supervised in face-to-face meetings with their faculty consultant. The pupils taught by the telephone supervised group did as well as the control group's pupils in learning about Landsat maps. The telephone group, however, reported much



dissatisfaction with telephone supervision. Since similar dissatisfaction regarding telephone conferencing was found at Open University, London as well, this must be considered as being a problem which has to be dealt with in future research.

Since telephone supervision was found to be associated with better or equal performance on two independent indices of student teacher effectiveness and since this method is more time-efficient and can be more economical, it is concluded that additional research on it should be carried out.

While the relationship depicted by the data permit several alternative explanations and hence cannot be regarded as providing unequivocal support for the claim that telephone supervision is more effective, or even for the lesser claim that it is *sufficiently effective*, the latter two claims are consistent with the data.

The study provides an empirical basis for pursuing further research on this method of supervision. In addition to probing the causal relationship between supervision and student teacher effectiveness, research should probe ways of making telephone supervision more palatable to the participants.

#### *Questions For Further Research*

Can more positive feelings toward telephone conferencing be generated by reducing the number of participants on line to allow more individual participation, limiting the time that each participant might speak, and providing thorough training for conference conductors?

Can one-way video used in conjunction with two-way telephone communication overcome feeling of impersonalness, by allowing the students to see their faculty consultant during the conference?

Is there a relationship between student teacher performance and the type of supervision by the faculty consultant?

Does a close relationship between the student teacher and cooperating teacher affect student teacher performance?

#### *Notes*

1. Edmonton Telephones, Meeting for discussion and demonstration of conference telephone bridging device, May 19, 1978.

#### *References*

- Carr, D. B. Teleteaching is inaugurated for students who are homebound. *California Education*, 1964, 1, 21-22.
- Dalrymple, J. I. Remote supervision at preservice and inservice levels in teacher education—Research reported in the vocational education special interest group at the American Educational Research Association meeting, 1971 AERA Convention. ERIC #ED050032 SP004900.
- Daniel, J. S., & Turok, B. Teaching by telephone, a two-way communication mode in distance education. *Proceedings of the 10th I.C.C.E. World Conference*, Brighton, England, May 1975, 133-140.
- Kirman, J. M. The use of infrared false color satellite images by grades 3, 4, and 5 pupils and teachers. *The Alberta Journal of Educational Research*, 1977, 23, 52-64.

## *Student Teacher Telephone Conferencing*

- Madden, C. F. Amplified telephone as a teaching medium; description of an inservice science seminar. In H. Bosley & H. E. Wigren (Eds.), *Television and related media in teacher education, some exemplary practices*. Baltimore, Maryland: Multi-State Education Project, 1967. ERIC #018978. EM006153. Article originally published in 1965.
- Smith, P. M. *Experimentation to determine the feasibility of remote supervision of student teachers*. Unpublished doctoral dissertation, University of Ohio, 1969.
- Stevenson, M. C., & Otto, A. C. Telephone teaching and testing: University of Nevada, Reno. *Journal of Home Economics*, 1975, 67, 27-28.
- Williams, E., Holloway, S., & Hammond, S. Students' reactions to tutoring by telephone in Britain's Open University. *Educational Technology*, 1975, 5, 42-46.

## PERSPECTIVES

### Some Observations Regarding Value Added By Education

Rupert N. Evans

*University of Illinois at Urbana-Champaign*

Although economists and tax experts have long been interested in assessing the value added by manufacture, little attention has been paid to the value added by education. The value added to an individual by an educational program can be expressed in terms of the individual's annual income, score on a job satisfaction scale, level of socioeconomic status, or any one or combination of measures intended to assess the effects of education. The value (however it be determined) of persons who are entering an educational program can be subtracted from their value upon completion of the program, and the difference is the value added by that particular educational program. Or, the value of persons who have not had a particular educational program can be subtracted from the value of a similar group of individuals who have completed a part or all of the program. Again, the difference is the amount of value added.

At first glance, value added by education may seem to be only another way of stating a concept which is usually referred to as "gains." Educational researchers have long used pre- and post-test designs which measure gains made by one or more groups which have received different treatments. Teachers have compared gains in achievement when they used different treatments. In almost all such cases, however, the consideration of "gains" is incidental. There has been an assumption that groups were nearly equal at the start of experimentation, and "gains" has been used almost solely for increasing the precision of measurement by making minor corrections in the starting point of similar groups. The ultimate criterion used in judging the education of an individual, however, is almost invariably the value of that individual as a finished product. This note, on the other hand, maintains that it is more important to consider the gains

---

Rupert N. Evans has served on the faculty of the University of Illinois (Urbana, Illinois, USA) since 1950, serving as Professor, Associate Dean, and Dean of the College of Education. Currently he is Professor of vocational and technical education in the Bureau of Educational Research. He holds the Ph.D. (in industrial psychology) and an honorary doctorate in vocational education from Purdue University. He has written many books and articles on work and education.



which are made by each individual. The sum of these gains is "value added by education."

At the present time, every phase of education except special education is concerned almost exclusively with value of the student as a finished product. I have noted that professors who work with doctoral candidates try to select those students who will reflect glory on their advisor and his/her institution. Vocational educators seek students who will be most successful in their careers. Music teachers seek students who are already talented musicians.

Tests used in education are usually concerned with value of the student as a final product, rather than with value added to that student by education. I have observed students who have learned to fail advanced placement tests so that they can take elementary courses and thus obtain a few high grades. This enables them to survive in a school which is less concerned with value added by education than with value of the final product. Teachers in schools which require standard examinations have learned to teach for the tests and to fight for above average classes in order to survive in a system that measures value of the finished product instead of value added by education. Teachers in every school I have observed prefer elective, rather than required courses, in part for the same reason. Athletic coaches search for talented athletes, rather than for students whose athletic abilities can be developed markedly. Their coaching abilities are measured by final results, rather than by value added. The dramatic coach, the band director, and the advisor of the school newspaper are under similar pressures.

The employment managers of manufacturing companies can and probably should have as their goal the selection of the best possible employees they can hire at the rates they can pay. The goal of most companies is increased value added by manufacture, and they choose people who can help to achieve this goal. On the other hand, it seems to me that the goal of educational institutions should be increased value added by education, so they should choose teachers and plan programs which can aid in achieving this end.

Unfortunately, at present, the ideal students are those who already know everything they will be taught. They are the students who will be the best finished products, and reflect the greatest credit on their teachers and their school. The ideal student is thus the one who does not need the education he/she will receive; the student for whom education will add no value. Something must be wrong with this logic.

Vocational education offers some of the most interesting examples of the need to measure the worth of education by the value it adds. Since the passage of the Smith-Hughes Act in the United States in 1917, vocational educators in that country have followed the precept that vocational education should be offered only to those people who can succeed in the occupation for which they are being prepared. There is nothing wrong with this precept. Obviously, if a person is trained to become a commercial artist, but he cannot succeed because no one discovered that he was color-blind, very little value has been added by the vocational education he has received. It is important that he be able to succeed in the occupation for which he has been prepared.

Unfortunately, many vocational educators all over the world have warped this precept. If they have a choice of students, they select those who are *most likely* to succeed in the course for which they are being selected or in the occupation for which they are being prepared, rather than choosing students who *can* succeed. For example, students who are amateur radio operators and who have designed and built their own electronics equipment are often preferred students in an elementary level vocational electronics course, even though they almost certainly will learn less in that course than someone who knows no electronics. The ideal apprentice often is one who previously has acquired all of the skills of a journeyman. In too many cases, the ideal student is one who does not need the educational program. He or she is one who probably can succeed without additional training.

It seems obvious that in many cases the person who will profit most from education is the person who has ability which is relatively undeveloped. If she has a handicap which can be removed easily by education, that education will have added great value to her and to society. If a very brief educational experience serves to identify and make visible a person who would not otherwise be considered employable, that education has added a great deal of value. If an educational experience develops ability for self-education, it is likely to add considerably more value than an experience which ends with the end of formal instruction. If an educational experience makes a difference between employability and non-employability, its value added is likely to be greater than the value added by an education which provides somewhat higher income to a person who will be employable regardless of added education.

In certain cases, however, the best candidates for training really are the best candidates who apply. A talented musician may become even more skillful under the guidance of a skilled teacher. A skilled automobile mechanic may be the best candidate for a course in diagnosis of automotive malfunctions. But in each case, the test of the selection of students and the worth of the educational program should be on the basis of the value which can be added by the training program, not on the basis of "value of the product."

We rarely are in a position to measure directly the values added by education. There have been attempts to measure the value added to earnings by government scholarships, but they have been contaminated by a suspicion that those who chose to use a scholarship were probably more likely to have succeeded with or without it than those who did not choose to use this educational benefit. Similar doubts cloud the many attempts to measure the value added by a high school or a college education.

There is, however, at least one clear case where data are available on value added by education. In 1962, Dr. Solomon Chaneles began a program of vocational training at Ruykers Island in New York City. Ruykers Island is a correctional institution for youthful offenders. When these young men were committed, they were given a test used by IBM for selection of computer programmers. A random half of those who passed the test were trained as programmers under a grant from the U.S. Department of Labor. The other half went through the standard education program of the institution. Both groups had the same raw material, but the value of the



finished product was clearly superior, according to every measurement, for the group which received vocational training. Even in this case, no precise measurement of value added by education has been attempted.

The precise measurement of value added by education is a problem which needs to be tackled by the educational economist, the educational psychologist, the educational philosopher, and the educational sociologist. All four groups need to be involved, because we should be concerned not only with the value added in terms of dollars, but also in terms of personal worth and worth to society.

Precise measurement of value added by education would be desirable, but is unlikely to come soon. In the meantime, however, much can be accomplished if the administrator, the teacher, and the public can begin to think in terms of value added by education. If they could simply ask themselves, "Will this action of mine add to the value of education for John Jones or Sally Smith", we would move forward much more rapidly. Entrance requirements to various educational institutions would undoubtedly be changed if we thought in these terms instead of considering only the value of the finished product. Nursery education might pay more attention to children from lower class homes. Scholarship requirements might be changed. Less attention might be given to tax rebates and more attention paid to scholarships based on potential for growth.

Probably the greatest difference would occur in the classroom. There is possibly considerably greater marginal utility in more attention by teachers to the poorest and to the best students in a class than there is in a concentration on the average student. If a student can be taught enough to keep him off relief for most of the rest of his life, the value added by education is probably far greater than a similar gain in knowledge by the average student. Similarly, a small amount of extra effort with brilliant students may lead them to self-instruction which will add a great deal to their value and to the value added by their education.



## ESSAY REVIEWS

### Essay Review I

#### EDUCATION IN ATLANTIC CANADA

D. A. MacIver

*University of New Brunswick*

EDUCATION AND DEVELOPMENT IN ATLANTIC CANADA. *Edited by Eric Ricker.* Halifax, N.S.: Department of Education, Dalhousie University, 1978, 381 pp.

There are no urban areas in Atlantic Canada that are able to absorb large numbers of rural immigrants. At least there are no centers to compare with Vancouver, Calgary, Toronto, Montreal, or even Saskatoon. These large cities are expanding urban areas which have been able to absorb much of the migrating rural population. As centers of power (and as the "consumers" of rural and other immigrant populations), cities have dominated the educational scene in North America regardless of the myth of the one-room school. The reasons for this development will not be argued here. It is the results of this process from the perspective of education and economic development in the non-urban environment that will be considered. The opportunity to do this is provided by the record of a conference, *Education and Development in Atlantic Canada*, which was edited by Eric Ricker of Dalhousie University.

This collection of papers brings together, probably for the first time, a variety of issues which reveal the problems of education and rural development. Unfortunately, it takes some effort to extract the significant material from Ricker's document and this is because of the way that Ricker has defined his task. When an editor undertakes to publish the record of a conference, he has to decide whether to be selective or comprehensive. Ricker chose the latter course and his collection of papers and discussions represents the good and the bad, the banal and the superb. This choice undermines the usefulness of the volume because some excellent contributions are hidden amongst a few tedious articles and some rambling and misdirected discussion. As a conference report this collection does its job. It is less successful as an analysis of the relationship between education and

---

Don MacIver has been a teacher, principal, university professor and Dean. He has edited two books on Canadian education and is currently concerned with the problems of non-urban education in Canada. Dr. MacIver has worked in Western, Central, and Eastern Canada, and has been Dean of the Faculty of Education, University of New Brunswick for the last seven years.

development in Atlantic Canada because the lack of selection disturbs the focus of the volume.

Having established these reservations, anyone interested in education and development in Atlantic Canada will be rewarded by slogging through the collection. Ricker's introductory article sets the stage and introduces a theme that is picked up in the better articles. The theme is that the schools in the Canadian East were never designed for the non-urban environment. From the 1920's on, there is evidence of a desire to centralize schools in order to provide an urban system of education. The blighted economic condition which exists in Atlantic Canada is due in no small part to the imposition of this urban concept of education. Unfortunately, it seems that neither the schools nor the universities have been able to provide an alternative that meets the needs of non-urban communities. Young people have drifted from the rural areas to the local urban communities which have never developed sufficiently to absorb all the rural population. The migration continues from the local centers to the cities of central Canada which have been forced to accept the people of the Atlantic region as immigrants.

The Atlantic Canadian has had a particularly bleak existence and the future is uncertain. Local entrepreneurs and multinationals have failed to establish a typical urban economy east of Montreal. So the Atlantic Canadian has been caught between a declining rural economy and an urban economy that has never recovered from a delicate infancy. These failures, so goes the argument, are compounded by the colonial mentality that pervades the East. This mentality has so undermined the confidence of the people that they are unable to see the potential that currently presents itself. Compared with urban communities, Atlantic Canada does not have to deal with major problems of pollution, crime, overcrowding and renewal projects. There seems to be a better future available, but habits of the past inhibit the imaginations.

Instead of seeking to capitalize on the indigenous resources, the politician and businessman in Atlantic Canada persist in trying to establish urban industries, much as an addled boxer reacts to a bell long after his fighting days are over. This is not unique, for almost nowhere in the world has education yet responded to the concept of non-urbanization (or re-ruralization) that seems to be in embryo in Atlantic Canada and elsewhere. More than one author wonders if the institution of education can respond to these new and sophisticated life styles. If teachers and professors remain as unresponsive as they have been in the past, then education is likely to become even less relevant to Atlantic Canada than it is at the moment. This is the general theme introduced by Ricker and pursued by the critical papers in this collection.

John Saul's papers considers the concept of Atlantic Canada as a colony of central Canada and other industrialized regions of North America. Schools "reinforce bourgeois domination and depersonalize the working classes," making them docile tools whose ambition is to be like the bourgeois (p. 74). There is the need to develop indigenous analyses and indigenous solutions to local problems, but the colonial tendency always to look elsewhere is encouraged by the educational system. Saul cites Tanzania as a country that is trying to protect itself from the "global ripoff" by



developing its own self-reliant economy. But this approach, Saul argues, is reflected in the system of education and there is a confidence in the people that seems to be missing in Atlantic Canada.

Saul's neo-Marxist critique seems to be accepted by David Macdonald. Macdonald is a harsh critic of the schools and universities. He sees them "as mindless purveyors of inappropriate goods, (and) our educational system has failed Atlantic Canada," but he goes on the claim that however witless, the schools remain the main hope. Again, like Saul, Macdonald cites the socialist Nyerere as a model leader, a man who is using education to build the confidence of his people, a confidence that the elite of Atlantic Canada do not have in their people. Macdonald accepts the apocalyptic theories of Illich and argues for a much less comprehensive educational system. While urging a diminution of the control of formal education, he argues for a much greater involvement of the community in education. One cannot help but think of Dewey's theories as one continues to read Macdonald. The paper gradually softens from the radical views of Illich to reflecting the later writing of Dewey in which the schools become continually more associated with the community. Clearly this is not a new idea, but perhaps its time has come.

Another of the more refreshing papers in the collection is written by Linden McIntyre, a veteran Halifax journalist. He argues that the system of education has been "criminally incomplete because it has not been responsive to local and regional situations" (p. 158). Again, one reads that the Maritimes have been treated as a colony and that industrial development has failed. What is needed is "small primary-based units . . . even if subsidized and protected by government" (p. 153). The theme that some third world countries have found some answers and that socialist policies may be useful is implied in this article as in the works of Macdonald and Saul.

If socialist ideas can improve the economic conditions of Atlantic Canada, it is not unreasonable to wonder about the absence of any significant radical party in Atlantic Canada. The interlude in this volume that deals with the Antigonish Movement suggests some answers to this question. Father Mifflin provides a careful description of this religious, yet broadly based social movement, and his comments are elaborated in the succeeding articles and discussions. The days of the Movement were exciting and stimulating even if "times were hard." The excitement did not always arise from overcoming the social conditions and fighting the bosses. Such activity occurred, but there seems to have been even more activity of a kind that was fatal to the development of a real socialist movement.

Sacouman argues that trade unionism (and by implication, socialism) did not develop in Nova Scotia due to the lack of "proletarianization." The typical "bluenoser" was a jack of all trades, a part-time miner, a part-time farmer, a part-time lumberman, and so on. When the pits closed he went back to the farm and did not join a labour union, but he was attracted to the Antigonish Movement which was a social force led by dynamic priests.

By the very nature of the leadership, the Movement played it cosy with authority. Coady, a founding father of Antigonish, was a friend of cabinet ministers and had a "harmonious relationship with the government" (p.



184). Indeed, Lotz points out that the government of the day assisted the Movement and as a result

It took a lot of weight off the back of the political system. If there had not been an Antigonish Movement, we might today be about twenty years ahead in terms of political reforms. (pp. 175-76)

The Antigonish Movement, it seems, undermined the development of socialism and other reform movements in Nova Scotia.

There were also problems within the Movement itself. The activist priests met with resistance from within the Church while the bishop of Antigonish, a supporter of the Movement, antagonized the bishop of Halifax, who was not. Then it seems that there was a special position adopted by the Scottish Catholics who became embroiled with the university . . . and so it went on. That, however, was only *within* the Church.

Social reform was undermined by religious differences. An occasion is cited when the United Church invited some representatives of the Antigonish Movement to a meeting. After a brief discussion all the Lutherans and Anglicans walked out leaving the Movement representative speaking to an almost empty hall. And then, of course, there were socialists and trade unionists whose efforts seem to have been emasculated by the Antigonish Movement. The Antigonish Movement overwhelmed the trade unions and socialist parties of the 30's and 40's and gradually the Movement diverted radical reformist tendencies into educational enterprises.

With a suddenness that jars, this fascinating discussion of the Antigonish Movement is dropped and the reader is confronted by a radical treatment of a contemporary issue. Edgar Friedenberg writes on the role of the schools in Atlantic Canada and his paper is crammed with insights. This paper is clearly and succinctly written, though one does not get that impression from the discussions that follow. Fortunately, it is not necessary to read the record of the discussion to get Friedenberg's point. In essence, he claims that the schools of Nova Scotia are the worst schools in Canada:

In the degree to which they discourage and make problematical for kids even the degree of self direction within the approved limits of choice which is essential for the most modern kind of acculturation, with all of its disadvantages. (p. 201)

Friedenberg repeats the assertion that Atlantic Canada is a colony and that schools train children to be colonized or to leave and head for central Canada. He emphasizes that there are various ways of perceiving the Maritimes and he enjoys the Nova Scotian lifestyle while recognizing its limitations. He has no desire to see Atlantic Canada become a "developed country," that is:

One that is in such command of its technical resources that it has the economic power to shift its industrial waste—human and otherwise—to elsewhere. (p. 213).

But there is an alternative that attracts Friedenberg and that is simply the idea of capitalizing on the local environment.

Friedenberg, for all his sophistication, is not so far from the views of many of the invited participants. They believe that education ought to have a good deal to do with improving the lives of people in the places where they live with their families and friends. Education ought to be concerned with many options and not merely that of becoming "developed," whether one

uses Friedenberg's idiosyncratic definition or not. Indeed, "development" is precisely *not* the question for Friedenberg. A common criticism becomes apparent. It is that few Atlantic Canadians have benefited from an education that has enabled them to participate fully in an urbanized and bureaucratized society. But most Atlantic Canadians are unable to see the possibilities of their environment because educators, functioning in response to the dominant pressures of society, are also unaware of the possibilities.

The theme of a misdirected approach to education is in Jim McNiven's paper. McNiven is an economist closely associated with industrial development in Atlantic Canada. This is not to argue that McNiven agrees with Friedenberg, because they take different positions and deal with different issues, but the fact remains that both believe that a more relevant approach to education is necessary in Atlantic Canada. McNiven argues that society has become overwhelmingly urbanized and bureaucratized which is not entirely a bad thing, even though society has been caught in the megamachine of bureaucratic control. He writes:

For most people throughout most of history, the adoption of these skills and values would have been a small price to pay for the stability and material standards that the megamachines offer. The advantages cannot be questioned. (p. 275)

With all its advantages, however, there is a move to "re-ruralization," that is, a movement of people back to the country because of dissatisfaction with urban existence.

McNiven takes great care to establish that the rural environment is not what it was. It is an environment that has been urbanized even though the population is sparse. That is, communication and organization have made isolated communities dependent on the contemporary social network. The school has been responsible for much of this kind of modernization and at the same time the centralized school has been responsible for the depopulation of the countryside. The structure of the school imposes an acceptance of a bureaucratic lifestyle which is another characteristic of urbanization. It is not possible to return to an idyllic past, but the schools do not prepare the young for the possibilities of a non-urban future.

As well as the lifestyle that one would anticipate, McNiven indicates that contemporary society has become a "single-crop society." He claims that if one large sector of our contemporary urban society were to collapse, the whole complex might settle about our ears. This contrasts with a rural society where, even under the most adverse conditions, whole societies may suffer calamities from drought and disasters of like kind and yet survive. McNiven's implication is that if we continue pursuing education as it exists today, disaster is inevitable. Only by seeking alternatives, especially rural alternatives, will it be possible to have some confidence that disaster may be averted.

A paper by Smitheram and another by Tilley reveal once again the urban orientation of schools but the perspective is different and McNiven's paper is nicely complemented, especially by Smitheram's work. Tilley's paper is a historical review of the decline of community control of schools in Nova Scotia. One of the interesting insights revealed in the paper is that before 1940 there was probably a lot less autonomy than one may have imagined.

Both of these papers contribute to a general picture of the decline of the rural school in Nova Scotia and Prince Edward Island.

The implication of this collection of papers is that the theory and practice of education in Eastern Canada has been urban-oriented for many years. This has probably been the case for education in the rest of Canada, but there is an important difference. That is, from Victoria to Montreal the process has worked *with* the prevailing economic conditions. In Eastern Canada the educational system seems to have been out of phase with the local environment for a long time. This has produced a colonial mentality and has oriented students into the wrong kinds of activities. Whether these claims can be justified in more careful and in more empirical studies than those contained in Ricker's collection is a moot point.

There are some difficulties with this volume which have been cited, but in large part it has served its purpose. To its credit, the work draws attention to the inadequacies of the schools and universities as they relate to the needs of people living in a non-urban environment. Essentially, this has been an urban century and education has tried to satisfy urban needs. The inadequacy of this approach, with related political implications, is fully revealed in Ricker's collection. Issues are raised, not settled; in doing this, the volume serves an important and useful function.



MORAL EDUCATION: THE RISK OF OVERSIMPLIFICATION

Ivan DeFaveri  
*The University of Alberta*

PREJUDICE. *Edited by L. Daniels, L. Douglas, C. Oliver, I. Wright, Association for Values Education and Research, University of British Columbia.* Toronto: The Ontario Institute for Studies in Education, Value Reasoning Series, 1978, 44 pp.; and accompanying TEACHER'S MANUAL, 32 pp.

THE ELDERLY. *Edited by L. Douglas, C. Oliver, I. Wright, Association for Values Education and Research, University of British Columbia.* Toronto: The Ontario Institute for Studies in Education, 1978, 40 pp.; and accompanying TEACHER'S MANUAL, 36 pp.

THE MORAL EDUCATION PROJECT (YEAR 5). *By C. Beck (Principal investigator), D. Boyd, E. Sullivan, J. Bradley, N. McCoy, S. Pagliuso.* Toronto: The Ontario Institute for Studies in Education, 1978, 180 pp.

Plato argued that the task of morally educating people was as difficult as, and inseparable from, the task of creating a just society. Many contemporary writers are now of the opinion that Plato made the problem more difficult than it really is, and we can now do with some ease what was once thought to be very difficult. We can now, it is alleged, by focusing on classroom activities—activities over which teachers have control, and without engaging in social and political reconstruction—do much more toward producing morally educated people than was hitherto thought possible, and this end result can be brought about without imposition, without manipulation and without indoctrination.

The task is complicated; the tendency to oversimplify must be resisted. Perhaps the most interesting oversimplification built into many of the contemporary programs of moral education is the claim that an adequate program of moral education need not concern itself with substantive principles because all that is really needed is the application of well-selected procedural principles. In short, it is sufficient to teach a methodology. This stance is a variation of the old slogan: do not teach students *what* to think; teach them *how* to think. Writers who follow this line of thought often invoke an analogy between the teaching of morality and the teaching of science: as it is possible to teach the scientific method without teaching specific scientific conclusions, so it is possible to teach a method appropriate to moral discourse while at the same time allowing the student to form his own conclusions. Moral education consists of teaching students how to be rational about morality, and surely no one could object to teaching students how to be rational. What could be more academically respectable than this? The task, then, is to teach a method and not to teach content, as the teaching of content would run too great a risk, there being no satisfactory way of deciding which content should be taught. The threat of indoc-

---

Dr. Ivan DeFaveri is Associate Professor in the Department of Educational Foundations at The University of Alberta. He received his Ph.D. at the University of California, Berkeley. His main interests are moral, aesthetic and religious values in education.

trinating or violating individual rights would be too great if content were allowed. The solution is not teach content at all. So it is argued.

I believe that this proceduralist account of moral education is seriously flawed, that a program of moral education can remain content-free only if it is simultaneously made trivial, and that any program of moral education that has the least importance must make reference to content. To insist on procedural principles alone is like saying to a student who wants to learn to play the piano that he should remember to play the right notes in the right order. In a sense, the advice is impeccable, but in isolation from some substantive comment about the correct order the notes are to take, it is of no significant value. Procedural principles that make no reference to content remain much too anaemic to form anything that could, without distortion of the language, be called a program of moral education. It is impossible to “do morality” in any important sense—to use an expression introduced by one of the writers in this school (Wilson, 1967, p. 26)—without doing a particular kind of morality, i.e., introducing substantive moral conclusions.

The programs that purport to restrict themselves to proceduralist principles rarely, if ever, exemplify the neutrality they advocate. Inevitably, they either smuggle in substantive values or, in various subtle and indirect ways, reinforce some of the values that are part of the social fabric in which they function.

Most of the English speaking writers in the area of moral education have been either English or American. We now have two Canadian organizations that have formulated programs of moral education designed for public schools. The Association of Values Education and Research (A.V.E.R.) at the University of British Columbia has recently published two small volumes, entitled *The Elderly* and *Prejudice*, and accompanying teacher's manuals. In Ontario, the Moral Education Project at the Ontario Institute for Studies in Education has recently published its final report describing the fifth year of its operation. In it, the reader is referred to publications by Beck for a more complete account of the program they are advocating.

The expressed aims of A.V.E.R. and the Moral Education Project are importantly similar. In the case of the former, we are told that their goal is to describe “how best to educate people in the domain of morality” (preface)<sup>1</sup>; in the case of the latter, “to help students reflect on their values . . . and . . . arrive at a sound set of values” (p. iii). But they have developed in individual ways and there are major differences between them. Both are at pains to preserve individual rights, to foster the autonomy of students and to avoid indoctrination. Both insist it is inappropriate to initiate students into any specific moral code. Each has made important contributions from which we can be guided in our attempts to influence the moral development of students, but both have serious limitation. Neither has entirely escaped variations of the proceduralist fallacy.

### *The Association of Values Education and Research*

A.V.E.R. tells us that any example of sound reasoning about values takes the same form. Regardless of how complex, they assure us, “its basic logical structure remains the same” (p. 8). It takes the form of a syllogism. The major premise gives the “value standard” or moral principle that is



pertinent. The minor premise gives the relevant fact or facts. The conclusion, which must follow logically from the premises before it can be accepted, gives the value judgment applicable to the case at hand. They give an example:

Value premise (principle)	It is undesirable to poison streams
Factual premise	Strip mining in the Kootenays will poison streams
Value Conclusion	Strip mining in the Kootenays is undesirable

(p. 8)

Their aim, “to help people to learn to be rational about matters of morality” (p. 9), must be seen in conjunction with their belief that one is rational about moral matters if one uses the practical syllogism well. Of course, rationality demands choosing only relevant facts in the minor premise and only an acceptable moral principle in the major premise. But what counts as a relevant fact? And what counts as an acceptable value premise? Let us consider each in turn.

Unfortunately they have taken only minimal steps toward answering the first question. They let their argument rest on two examples, but it is impossible to extract from them any principle which may serve as an answer to the question.

The first example deals with a technical matter in elementary mechanics: “If I want to keep a large, heavy door from swinging shut but do not have a door stop,” and I have a choice between a light pillow and a heavy stone, I would be irrational to choose the pillow (p. 4). Fair enough. If I need something heavy, and I have a choice between something heavy and something light, I should choose the heavy object. But whatever principle can be extracted from this example will not guide us in any important way if we are dealing with a genuine moral problem.

The second example describes a personnel manager who must hire someone for a position, the sole qualifications for which are competence and industriousness (p. 4). We are told that under these conditions, it would be irrational to consider the racial origin of the applicants. Of course. The previous decision that the “sole” qualifications are competence and industriousness make this conclusion inevitable—and uninteresting. But those people who find questions about race troublesome never have problems quite like this. In the example, the question whether race is relevant is not itself problematic, but it is precisely the question of whether or not race is relevant that defines the actual problems in some actual situations. Today, for example, the advocates and the opponents of affirmative action programs differ over precisely this point. Similar arguments abound in regard to the relevance of age in, for example, the amount of insurance premiums that a person ought to pay. Even in the example given, it is easy to imagine a different personnel manager, one who had yet to decide on the qualifications for the job, wondering if ability to get along with existing employees is an appropriate qualification, knowing that these employees were racial bigots. Here, he might on the one hand want to take advantage of the opportunity of civilizing them and make a point of



hiring someone from a different race, or on the other hand he might want to avoid the risk of decreasing their morale and lowering their productivity.

These examples tell us that what facts are relevant in any genuinely controversial matter is often itself a matter of dispute. Sometimes it is precisely this which distinguishes rival moral opinions. Insofar as the writers have raised the question of relevancy but have appealed only to unhelpful examples in their attempt to answer it, they have really avoided the question. I'm afraid some students may get the impression that it is easy in principle to decide which facts are relevant and thereby be left with the conclusion that making moral judgments is simpler than it really is.

Difficulties in deciding on the appropriate factual premise are of course paralleled by difficulties in choosing an appropriate value premise. In regard to this, the discussion is more detailed. They provide us with four tests designed to distinguish a major premise which is rationally acceptable from one which isn't. Any value premise which passes these tests must be seen as being as good as any other which passes the same tests. The tests, in summarized form, are:

1. *Role exchange test*: Ask if you would be willing to exchange places with the person(s) most disadvantaged in a situation. Imagine yourself in the circumstances of the most adversely affected person or group, then decide to accept or reject the value judgment and its action in the light of the consequences to the most adversely affected (p. 10).
2. *Universal consequences test*: Ask if you would be willing to accept a state of affairs where everyone acted according to the same principle (p. 10).
3. *New cases test*: Ask if you can accept all the judgments that issue from the principle in question. If there are legitimate exceptions, they should be built into a new principle (p. 11).
4. *Subsumption test*: Ask if your value principle follows logically (can be deduced) from a higher order value principle which is acceptable to you. If you cannot accept the higher-order principle, find new reasons for holding your original value principle, or reject it and reconsider your value judgment (pp. 11-13).

It is worthwhile noting just what these four tests accomplish, and what they do not. They are addressed to the person who has the motivation to be moral, but who wishes to be sure that the major premise of his practical syllogism is rationally justified. The tests, we are told, are for people who want to be "rational" about matters of morality, and who want to understand the "structure" and "logic" of morality (p. 8).

These tests have nothing to say to the person who sincerely wonders why he should enter the domain of rational morality in the first place, or who has no inclination to do so. They do not address themselves to the person whose motivation is such that the question of justification does not arise for him. Nothing here, in other words, is pertinent to the case of the egoist, who may understand the logic and the structure of morality as well as the next person, but who happens not to care about the well-being of anyone other than himself, and is quite willing to let others be egoists too. Similarly, the egoist may "imagine" the consequences for "the most adversely affected" people just as effectively as the next person, but if he doesn't care about the well-being of others, he may accept a principle that others are certain ought to be rejected. In light of the self-assigned task to see "how best to educate people in the domain of morality," it is difficult to see why they are

concentrating solely on the logic of justification and have so little to say about the psychology of people who may or may not want to use this logic. In the education of human beings, the education of the emotions cannot be so easily dismissed.

Let us leave the egoist and assume we are dealing with people who have the appropriate motives; that is, people who want to be moral, but don't quite know which "value premise" is morally acceptable. How much guidance would these four tests give him? Buddha's life changed, according to one legend, as a result of his imagining himself in the place of the aged and infirm. Many people, as a result of applying these tests consistently, would be persuaded to accept one value premise rather than another, and the world would probably be a more civil place if more people did apply them. People often judge their own conduct by a different set of rules from those by which they judge the conduct of others, and these tests are designed to thwart this tendency. But to say that many people, as a result of applying these tests and the recommendations implicit in them, will be *psychologically persuaded* to follow one path rather than another, is different from saying that a person applying them will be *logically compelled* to come to a particular conclusion. Given certain psychological dispositions, these tests would tend to rule out certain major premises and encourage others, but given a different set of dispositions, they could very well be used to support premises ordinarily found reprehensible. One can love others as much as he loves himself, but not love himself very much. The extent to which these tests assist in a program of moral education depends to a large extent on the kind of person who applies them, and how he describes the situation he is confronting.

Consider the role exchange test. Under certain conditions it could be seen as a very strong test, under other conditions very weak. If you took it very seriously, and in isolation from all considerations of desert or utility, it could be used to support a radical egalitarianism in which no one's total well-being was any worse than anyone else's. It could be used to argue that since no one, all else being equal, is willing to exchange places with anyone less well off than he is, then we ought to have a society where everyone's total good is identical. Only such a society would be a moral one. The differences tolerated by non-egoists are precisely those differences which they allow as a result of an appeal to their principle of justice, or their principle of utility, or because they believe such differences appropriately reflect innate differences. Without an appeal to any such putative justification for such differences, a radical egalitarianism such as this would be more rational than allowing the differences to continue. Viewed in this light, advocacy of this test can be seen as tantamount to advocacy of a particular substantive moral theory, presumably something the authors want to avoid.

But those who favour such a test rarely reach the conclusion that this kind of radical egalitarianism is appropriate. This is because it is rarely, if ever, used in isolation; it is almost always used in conjunction with principles of justice or utility, which it complements but does not supplant. If it is applied while simultaneously respecting certain moral principles having to do with desert or usefulness, it may alter nothing. There is a sense in which a person who restricts himself to moral considerations, and ignores



all other considerations such as those of prudence, aesthetics and religion, must be willing to exchange places with anyone who is treated morally. It is to this I now turn.

To say of someone that he is “adversely affected” or “disadvantaged” is to make a comment about the differences that exist between him and others. But these differences may be seen as morally acceptable or unacceptable. A person may lack certain material goods or opportunities, but it may be felt that his having them would not serve the common good or his having them is in some way deserved. In a hierarchical society such as Plato envisioned, for example, justice demanded that some have more than others. Usually, however, the expressions carry the suggestion that the disadvantaged person is unjustly disadvantaged and that the situation needs to be improved. Imagine a person who believes existing differences satisfy his sense of justice or his principle of utility, or who believes that in some way they are morally appropriate. He may believe, for example, that a person’s well-being ought to be in direct proportion to his contribution to society. This person may be willing to consider the most adversely affected person, and to consider the possibility of exchanging places with him, but not without simultaneously asking whether each deserves his respective position or not. Of the person who is most adversely affected, he might say: if I contributed as little as he did, I would deserve as little as he is getting, but to receive less than I am now receiving while simultaneously continuing to make my present contribution would indeed be unjust. Nor would I exchange places with someone who is better off than I am if he deserves his better position; that also would be unjust.

It can be seen, therefore, that if a person has been treated justly, he cannot be said to have been disadvantaged in any moral sense. If his treatment has been morally satisfactory he can continue to be “disadvantaged” only in a nonmoral sense, with which this program is not concerned. Of such a person our hypothetical speaker may say that although he may not want to exchange places with him on, say, aesthetic grounds, he would be willing to change places with him on moral grounds, because both are being treated in a morally correct manner, and no one has the right to expect more than just treatment. Thus, such a person, that is, one who believes existing differences are justified would, if he restricted himself to normal considerations, be willing to change places with anyone. But it should be noted that his willingness to do this may be combined with a view of justice we find offensive. Thus the test may not accomplish what the advocates of moral education hope.

A second reason why this test may not do as much as its advocates hope becomes obvious once we realize that two people, both of whom may be willing to apply the test, may be unable to agree as to who actually is the most adversely affected person in a particular situation. To describe someone this way is not to offer a neutral description, but is almost always to offer a description from within a particular moral code. A person who holds a moral code different from ours may isolate a different person from the one we do, even though we both use the same words to describe our respective candidates. Thus, someone with a bizarre moral theory may acknowledge the legitimacy of the Role Exchange test, but his doing so



would not in any way improve the adequacy of his moral theory.

The status of the Universal Consequences test is in many ways similar. Many people will find in it a psychological persuasive reason to do one thing rather than another, but others won't. The test itself offers no logically compelling reason why students should accept any particular substantive conclusion. Some students may well appeal to this test as a way of legitimizing a conclusion others find repugnant, and do so on the grounds that since they are willing to have everyone follow the same principle, the principle ought to be tolerated. The example given as an illustration of the use of this test will make this clearer. Ted has cut across the grass:

Ned: You shouldn't do that.

Ted: Why not?

Ned: Because you will wreck the grass.

Ted: Phooey. How could I do any serious damage just by walking across it once?

Ned: That may be true, but what if everybody walked across it?

Ted: W-e-l-l, I suppose that would wreck it. But so what?

Ned: It is easy to see what. If everybody walked across it, it would wreck the grass. And everybody should be treated fairly. That means we should all be treated alike unless we can show that there is something special about us (or our circumstances) that merits our being given special privileges. So unless you can show that you are special in a relevant way, you should stay off the grass. (p. 10)

Even in this simple example, one can see that this test carries weight only if there is previous agreement in regard to several substantive matters. There is no argument against the person who sincerely believes that the advantages of having a short cut outweigh the disadvantages of ruining the grass.

The implication left by the advocacy of this test is that any value principle that one is willing to universalize is morally acceptable. Now often this test will be sufficient to convince some people to modify their behaviour: they may realize that their actions should be judged by the same standards as they judge the actions of others, and thereby modify the behaviour that was about to make of them an unjustified exception. But in many cases of genuine moral dispute, it will have no force. Both sides of the abortion dispute are willing to universalize the principle that they see operative, but happen not to agree on what that principle is. Both are willing to universalize the principle that one ought not to murder, but the disagreement is about the metaphysical issue as to what counts as a human being. In regard to the same act, one would describe it as a case of murder and the other would not. The universal consequences test will not resolve the dispute, as both sides are willing to appeal to it. A pacifist and a militarist are both willing to live in a world in which everyone did what they were doing, but this in no way diminishes the differences between them. Egoists, mercenaries, amoralists, and fanatics are usually quite willing to have everyone follow the principles by which they live; often they are quite unhappy precisely because others don't.

The principle is effective if one who does not believe he has the right to think of himself as an exception finds himself doing just that. It will remind him that he ought to subject himself to the same rule he applies to others. Further, it may bring about a certain commonality of response among people who describe a moral problem the same way and who share a prior

commitment to certain substantive values. But it has no force against the person who is quite happy to see himself as unique and who applies unique standards to himself. Further, this principle, even if applied by everyone, is unlikely to increase agreement among people who sincerely describe moral problems differently.

In actuality these two tests, even when combined with the remaining two, do nothing but demand of the moral agent a pervasive consistency: he must judge identical acts the same way, he must prescribe to himself the same principle he prescribes to others, and the practical syllogisms he invokes must be internally consistent. But I think a person could successfully subject his major premises to each such test and still be the kind of person we would regard as morally miseducated. This "logic" in the hands of ordinary, decent people would allow them to be moral more consistently; but in the hands of the egoists, or the selfish, or those with odd emotional lives, or those with distorted theories of justice, it would be quite impotent. But it is precisely with these people that a program of moral education should be most concerned.

Throughout their discussion about "the logic of value reasoning," a position is taken in regard to the fact-value controversy, which is perhaps the issue more discussed in 20th century moral philosophy than any other. The position taken by A.V.E.R. is reflected in the centrality given to the practical syllogism and the process of deductive logic embedded in it. The message that the student will receive is that there is always a clear, unambiguous and easily defined difference between facts and values; that is, there is always a radical difference in kind between facts and values. An exercise given to the students will illustrate this. From a series of statements, students are asked to select those which are factual claims and those which are value judgments. For example, the statement, "If an ethnic group is unable to retain its language, then it loses its culture" is, we are told in the *Teacher's Manual*, a fact. But it would be just as correct to say, given the information that the student has, that this statement is nothing more than a definitional commitment, or that this statement emerges as a "fact" only if certain values are agreed upon, as the notion of "loss" and indeed the notion of "culture" are value-laden. Whether or not a sharp distinction can be made between normative and descriptive disciplines is deeply problematic, but the authors have settled for a particularly heavy-footed distinction, one deeply rooted in positivist thought throughout the centuries. This view has appeared, faded, and reappeared throughout history. I do not want to argue for or against the position taken by A.V.E.R. about this controversial matter, or whether Schopenhauer is right when he calls the position taken here "the greatest and most pernicious of errors, the fundamental error, the intrinsically *perverse* view" (1970, p. 133, italics in original). But I want to point out that I believe it is suspect, from the point of view of education, to pass off without discussion of its logical status, something which is really problematic. We are told that "this program proceeds on the assumption that there is a rational approach to the solution of moral problems" (p. 5). Well and good; but the further, unstated assumption is that there is only one rational solution to the fact-value problem. Some students will not recognize this as a further assumption; indeed, they will believe it is a "fact" that their world is divided into two



discrete and mutually exclusive categories, and that value-free nature contrasts sharply with man's imposed evaluations.

### *The Moral Education Project*

The publication of the Moral Education Project at O.I.S.E. here considered describes the fifth year of their operation. It begins with a summary of the Project's "reflective approach" to values education, the justification of which is given in greater detail in other publications (specifically, Beck, 1971, 1972, 1974, 1976). The program is also referred to as the "fundamental human values" or "ultimate life goals" approach because of the pivotal place ultimate values are seen to occupy. The function of a moral education program is to show people the best way to attain their ultimate values. Of ultimate values themselves, three claims are made. First, human beings tend to pursue the same ultimate values. They tell us that "as we become aware of the connections between our life goals, we will notice that there is more in common amongst the life goals of people around the world than we had previously thought" (Beck, 1972, p. 10). Great commonality among human beings in regard to ultimate goals is envisioned:

The great majority of human beings pursue ultimate life goals, or fundamental human values such as survival, happiness (enjoyment, pleasure, etc.), health, fellowship, (friendship, love, etc.), helping others (to some extent), wisdom, fulfillment (of our capacities), freedom, self-respect, respects from others, a sense of meaning in life, and so on. (Beck, 1976, p. 1)

Second, each ultimate life goal must be given equal status. "The old notion of a hierarchy of values . . . is quite unworkable . . . . It is not possible to develop a ranking of goals such that one can simply read off which to pursue when two goals conflict" (Beck, 1976, p. 269). If they do conflict, he tells us, unhelpfully, that, "appeal must be made to the maximization of life goals, in the long run and on the whole and all things considered" (Ibid, p. 269). Third, ultimate life goals cannot be subjected to rational scrutiny. People "just *do* pursue them" (Beck, 1974, p. 271, *italics his*). Rationality enters only as a judgment on the means chosen to attain these ends, but not as a judgment on the ends themselves. He is quite adamant about this. It is "meaningless", he insists, to ask whether people in general ought to pursue the goals they actually do pursue (Beck, 1974, p. 24). Elsewhere we are told that he is talking about "mature" human beings only, and it is necessary to reject the choices made out of ignorance, prejudice and superstition (Ibid, p. 25). There are a number of problems with this view of the valuing process. To say that the great majority of human beings tend to want some or all of the ultimate life goals listed here is misleading. At this level of abstraction, indeed almost all people will tend to use the same words to describe what they are after, but the agreement may be purely verbal. It is all very well to say that both Arabs and Israelis want justice—indeed they do—but the sense in which they want the same thing can be seen as trivial, especially at the onset of each war. Both the Marquis de Sade and his victims wanted happiness. Both capitalists and socialists strive for the fulfillment of human beings. Such terms serve a purpose, but in isolation they do not tell us anything important. Only if the context is made clear, or if the statement of the ultimate goal is followed by a series of progressively more specific



statements about the goal, will it be sufficiently informative in any reasoned account of how we should govern our moral lives.

The claim that people just do accept certain goals as ultimate, and that the task of moral education is merely to give them the ability to attain these goals efficiently is a philosophically respectable position, but there are other conceptions about the nature of ultimate goals that need to be taken seriously and about which a teacher genuinely interested in education would insist on his students knowing.

The orientation reflected here is embedded in the empiricist tradition and is sometimes encapsulated in the slogan, “good reasons apply to means, not ends.” Any end is morally acceptable as long as attaining it does not interfere with the pursuit of a more valued end. The ancient tension between what men “just do” in regard to the choosing of ultimate ends, and what they should do—either because of their own nature, or because of the structure of the world—dissolves here. This view, embedded in the ultimate life goals approach, contrasts with various intellectual traditions which claim that even ends can be judged rationally, and that man, not entirely immanent in time, can in some important sense transcend himself. The former view sees morality as an applied social science and the making of moral decisions essentially as a technical matter. The latter view believes that morality, properly conceived, reflects an independent, critical, and autonomous reason and rejects the view that reason is purely instrumental. This continuing debate in intellectual history and philosophy is far from over. As Kolakowski reminds us, “philosophy has never given up its attempt to constitute an autonomous ‘reason’ independent of technological applications” (1969, p. 210). But this highly problematic matter is here treated as if there were no dispute; insofar as this is done, the purposes of moral education are poorly served. If we are interested in educating students in the domain of morality, rather than just encouraging them to adopt a particular moral theory, we should make sure that they are familiar with the most important controversies that are part of its evolving history.

Even with the conception of ends that is entertained here, it is still odd to say that people “just do” choose ultimate ends, especially in an educational context. For people choose one end rather than another in part because of certain experiences they have had and certain influences that are exerted upon them. If teachers don’t encourage students to choose one set of goals rather than another, it will be done by default by advertisers and con men of various persuasions. Further, the list of fundamental life goals given here reflects a curiously optimistic view about the nature of man. It is as if all his fundamental impulses were good, and there were no darkness in his heart. But many people, I suspect, if they were not self-deceived and if social conditions allowed, would truthfully tell us that a list of their ultimate goals would include such things as self-aggrandizement, mastery over others, and a sadistic impulse to destroy. It is on the periphery of this point that Beck makes a startling admission, but one the significance of which he seems to be unaware. In regard to the “reflective approach,” he says, “In order to reflect successfully on values, a student must already have arrived at a relatively sound value outlook” (1970, p. 3). It seems that the approach only works with students who need it least, and nothing has been said about

what is to be done with those students who need moral education most. In our school system, it is not the student who has already arrived at “a relatively sound value outlook” who is bothering people; it is the student who is the egoist, or who is insensitive to the needs and feelings of others, or who is not influenced by moral arguments.

As was suggested earlier, no attempt has been made to evolve a principle which can settle conflicts, either conflicts which arise because a person holds mutually exclusive ends, or one which arises between two people. We are told only that these people “should be shown that the major concern of ethics is to try to arrive at mutually beneficial arrangements or ‘happy compromises’ ” (Beck, 1971, p. 33). This is advice which every high school girl with an unwanted pregnancy should read very carefully.

### *Conclusion*

The problems that arise when an attempt is made to initiate a program of moral education are very difficult in a country which celebrates cultural diversity, respects individual rights and personal autonomy, believes education is something beyond socialization, and insists on compulsory education. Perhaps it is impossible to have a program of moral education without modifying some of the good things simultaneously wanted. The two Canadian programs represent bold attempts to satisfy the considerable demand for practical programs in this sensitive area. Insofar as they will encourage students to realize that it is often inappropriate to hold values at a pre-critical and pre-reflective level, and that it is often necessary to display public reasons for one’s moral conclusions, they will have done a great service. But this promise of future success must be tempered with an understanding of the limitations the programs themselves impose.

In regard to A.V.E.R., one is struck with the incongruity between their reluctance to enter the area of content and the forthrightness with which they commit themselves to a particular, controversial view of moral reasoning. As I have argued, I believe the first step here is to realize that a program of moral education which restricts itself to procedural principles alone will accomplish less than what most people expect from it. It is to the credit of the Moral Education Project at O.I.S.E. that they have seen the need to discuss substantive values. But in their assumption that, after all, everyone’s goals are similar and that being moral is nothing more than the efficient attainment of these goals, they have treated a long-standing dispute in the history of thought as if it had been finally settled. In this regard, perhaps neither program takes the notion of education sufficiently seriously. Students guided by such programs could quite easily be misled about the status of certain claims made by the authors. They might not realize that the positions taken in regard to, for example, the nature of morality, the scope of deductive logic within moral discourse, the adequacy of the means-end model in this area, and the nature of rationality, are themselves quite controversial, and by no means the only positions entertained by serious writers in moral philosophy. I find this an unhappy situation, because I believe that although a teacher in a public school need not be impartial in regard to controversial matters, he ought to represent competing positions in a fair manner. By not telling the students about



alternative accounts, the writers are being less than attentive to the desirability of promoting autonomous students, an ideal to which they implicitly subscribe.

Thus, these programs restrict themselves to a limited series of questions. A more complete investigation of what it is to be morally educated would include a study of the forces which help shape what one of these programs refers to as our ultimate goals and the other program as our willingness to live by certain principles. This would incorporate, at least, a study of insights from psychology, history, and sociology. All the programs mentioned in this essay reflect an individualist stance. But if we believe that, in part, a person's self is an historical and social product, we will not be satisfied with a program of moral education until it considers the possibility of influencing the factors which help form the individuals involved. Thereby, we will realize that the problem of creating moral people must be seen, at least in part, as the problem of social justice. Similarly, a program of moral education ought not to overestimate the value of instrumental reason. Perhaps rationality, in this sense, is not all we need. G. J. Warnock has written that if people choose not to want to be moral, "the essence of the remedy is not in reason, it is in non-indifference" (1973, p. 166). One can "reason" and "reflect" but be insensitive to the suffering of others. What is it that will bring the capacity to care about others into being, and make a person realize that the pain he himself suffers is no more important than the pain that others suffer? An instrumental view of reason will not help here, because according to this conception, the notion of reasonableness arises in the area of morality only if such a commitment is presupposed; to decide to care about others would be a pre-rational or nonrational commitment. But if one doesn't care about others, neither talk about "logic" nor a systematic listing of actual desires will do much good. Some, as was noted earlier, will really want to pursue destructive ends regardless of what they profess. We may discover that in the final analysis, being moral involves more than having mastered a technique.

#### Notes

1. All questions from the A.V.E.R. material are in *Prejudice: Teacher's Manual*. The same quotations appear in *The Elderly: Teacher's Manual*, but the pagination is different.

#### References

- Association for Values Education and Research. *Prejudice*, ed. by L. Daniels, L. Douglas, C. Oliver, I. Wright. Value Reasoning Series, The Ontario Institute for Studies in Education, 1978, and accompanying *Teacher's Manual*.
- Association for Values Education and Research. *The elderly*, ed. by L. Douglas, C. Oliver, & I. Wright. Value Reasoning Series, The Ontario Institute for Studies in Education, 1978, and accompanying *Teacher's Manual*.
- Beck, C. *Moral education in the schools*. Toronto: The Ontario Institute for Studies in Education, 1971.
- Beck, C. *Ethics*. Toronto: McGraw-Hill, 1972.
- Beck, C. *Educational philosophy and theory*. Boston: Little, Brown and Company, 1974.
- Beck, C. *The reflective approach in values education*. Toronto: The Ontario Institute for Studies in Education, 1976.
- Kolakowski, L. *The alienation of reason: A history of positivist thought*. New York: Doubleday, 1969.



- The Moral Education Project (Year 5)*, C. Beck (Principal Investigator), D. Boyd, E. Sullivan, J. Bradley, N. McCoy, S. Pagliuso. Ontario Ministry of Education, The Ontario Institute for Studies in Education, 1978.
- Schopenhauer, A. *Essays and aphorisms*. Harmondsworth, Middlesex: Penguin, 1970.
- Warnock, G. J. *The object of morality*. London: Methuen, 1973.
- Wilson, J. *Introduction to moral education*. Harmondsworth, Middlesex: Penguin, 1967.

## BOOK REVIEWS

CLASSROOM LANGUAGE: WHAT SORT? *By Jill Richards*. Winchester, Mass.: Allen & Unwin Inc., 1978, 154 pp.

*Classroom Language: What Sort?* explores issues related to language learning and learning through language in schools. Professor Richards introduces this exploration with reference to what she identifies as a current "dilemma for schools." This dilemma is seen as the result of pressures to educate for the demands of modern technological society coupled with the concern for providing opportunities for personal growth and development. The implication of these competing demands for classroom language are discussed with reference to present practice and to the thinking and research of prominent educators and linguists. Professor Richards' own research in British schools provides input to the six-chapter text which challenges some popular beliefs related to school practice and urges some rethinking about the nature and scope of school language programs.

If there is a fundamental argument guiding the discussion in the text, it is a plea to resist simple answers or solutions in treating such complex phenomena as language learning and language itself. What are described as "language-limiting" practices and "class" stereotyping of students are cited as prime examples of simple and inadequate responses to complex problems and issues. It is argued, for example, that many so-called "child-centered" approaches to language learning in the early school years are, in fact, inadequate with respect to language function emphasis and too restrictive in terms of the experiences provided for students. Various practices associated with developing "self-expression" and "creativity" are cited as examples of focusing on "poetic function" to the exclusion of other language functions, and emphasizing "first-hand" experiences to the exclusion of other viable alternatives in providing opportunities to learn and use language.

With respect to the language learner or "user," Richards argues against the use of social class stereotypes in explaining and predicting language behaviour. Stating that much of the sociolinguistic research has produced conflicting conclusions, she reviews the work of Bernstein ("codes") and Labov ("dialect") in particular. Again her argument rests on the criterion of adequacy. She claims that the theories of language learning and development restricted to social class influences are inadequate explanations of complex phenomena. She argues further that we will come no closer to understanding language behaviour until we look at a number of factors influencing language within individual, rather than group, life-style. She

notes, also, that many children whose language does not represent “code” or “dialect” have difficulty meeting some of the language demands placed on them. Richards states that this is particularly evident in what is described as the transition from primary to secondary school.<sup>1</sup>

The kinds of problems associated with the language and thinking demands of programs in the later school years are identified in discussions of “language demands of subject learning” and “specialized language and concept development.” Although some of the difficulties experienced by older students are traced to the “language-limiting” nature of primary level programs, Richards does not excuse subject area teachers from any responsibility in this regard. Much of the discussion related to language demands in subject areas centers on Richards’ research in classrooms. The research was designed to test hypotheses related to children’s “ordinary language” and the language encountered in school, different language demands of different subject areas, and the linguistic features of the language in different subject areas. The reports of Richards’ investigation and the work of other researchers identify the conceptual tools and systems of analysis used in this kind of research on classroom language. The findings of this research, and the discussion of specialized language and concept development related to subject learning, raise questions which have implications for teachers’ language and instructional strategies and techniques. The main thrust of these discussions appears to be an attempt to develop awareness of the role of language in learning across the curriculum. Implicit also is the notion that the school language program needs to encompass all facets of school life and learning.

The “conclusions and recommendations” set forth in the final chapter suggest how teachers might make use of the text’s discussions in examining language program and practice. Pertinent recommendations from the Bullock Report (1974) are reviewed also in relation to issues raised throughout the text.

Professor Richards has identified and analyzed a number of significant issues in a relatively slim volume of 144 pages. Although there is an attempt to link discussion throughout the text, chapters 2, 3, 4, and 5 are, in a sense, self-contained with respect to focus and argument. Thus the text is best described, perhaps, as a series of discussions on different issues in language education. However, this feature of the text suggests its value as a reference or resource work. With one exception, the arguments in the text are clearly developed and not overstated. The exception is the attempt to challenge the conclusions of some sociolinguistic studies. In this discussion and analysis, Professor Richards comes close to arguing herself into a “semantic” corner. Nevertheless, the text stands as a scholarly piece of work which should be of interest to teachers and to all concerned with language programs in schools.

#### *Notes*

1. The designations of “primary” and “secondary” school refer to the organizational structure in England. Children of 12 years are identified as being in the first years of secondary school.

Ellen M. Regan  
*Department of Curriculum*  
*The Ontario Institute for Studies in Education*



AN EXPERIMENT IN TEACHER EDUCATION. *By Alec Ross, David McNamara and Jean Whittaker.* Guildford, Surrey: Society for Research into Higher Education, Ltd., 1977, 133 pp. (ISBN 0-900868-554)

This monograph is based upon the final report of an experiment in teacher training which was undertaken in the northwest of England in the period 1971-1974. It involved some fifty-two students in three small colleges associated with the University of Lancaster. The main thrust of the experiment was to substitute courses, in which the subject matter of several disciplines taught in schools was "applied" to the education of children, for courses of considerable depth in one or two academic disciplines. This concept is in line with the thinking of the Bicentennial Commission on Education for the Profession of Teaching which completed its work in the United States in 1976. This Commission recommended that principles be the "constants" of academic study rather than have students concentrate on a maze of specific topics. The logic behind this substitution of courses, or, in the American context the emphasis on principles rather than on the details of a discipline, is straightforward. It is that teacher education should be designed to help teachers teach youngsters the basic elements of the disciplines rather than to aspire to make prospective teachers into mathematicians, scientists or whatever.

The actual "Applied Education Course" attempted to cover the subject content appropriate to the learning of children of the 5-13 age group. Modules concerned with key areas of the schools' curricula were developed. Two modules, English and Mathematical Education, were compulsory with students given choices among several others. The overall Applied Education package involved considerably more than merely the teaching of subject-matter but included the development of attitudes and skills. These additional objectives may have had some confounding effects on the experiment.

The description of the experiment is exceptionally detailed and rather elaborately presented so that the reader tends to get an exaggerated notion of its significance or sophistication. The principle involved is an important one, but a relatively small sample, together with a rather hurriedly conceived and perhaps amateurishly implemented test, should not persuade too many to rush to follow suit. The findings, for what they are worth, do seem to indicate that such an approach tends to be better preparation for teaching in "primary" schools than is the traditional in-depth subject matter route. College staff appeared to welcome it and students were judged to be more enthusiastic. There was no indication that such a course failed to "stretch" intellectually the experimental group. In fact, those of the experimental group who continued their professional education to the B.Ed. level did particularly well.

The general impression conveyed by this account of the experiment is that, for students firmly committed to careers in primary teaching, study of the application of discipline content to teaching is satisfying and appropriate. For students lacking this firm commitment such an approach has less value. On balance it appears to this reviewer that the monograph is a useful addition to the literature concerned with teacher education but that its substance could have been conveyed in considerably fewer pages.

Lorne D. Stewart  
*Department of Secondary Education*  
*The University of Alberta*

# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

## *Volume XXV, 1979—Table of Contents*

- Adams, B. B., and Ollila, L. O. *The relationships of language concepts, auditory comprehension, visual perception, and spatial relations as predictors to reading achievement in first grade*. No. 4, December, 248-258.
- Aikenhead, G. S. *Using qualitative data in formative evaluation*. No. 2, June, 117-129.
- Assheton-Smith, M., and Toohey, K. *School-centered community conflict: The Holdeman Mennonite case in Alberta*. No. 2, June, 77-88.
- Bacchus, M. K. *Education as a social control mechanism*. No. 3, September, 160-173.
- Bragman, R., and Hardy, R. *Identical and reverse visual pattern recognition in deaf children*. No. 3, September, 174-181.
- DeFaveri, I. *Moral education: The risk of oversimplification* (Essay Review). No. 4, December, 294-306.
- Evans, R. N. *Some observations regarding value added by education* (Perspective). No. 4, December, 284-287.
- Hamalian, A. *Pupil control ideology: Comparative Perspectives—United States and Canada*. No. 1, March, 37-47.
- Horodezky, B. *Comparative difficulty of beginning reading vocabulary: Set II*. No. 4, December, 259-263.
- Hughes, A. S. *Curricular knowledge organization and variations in instructional emphases*. No. 1, March, 4-19.
- Kehoe, J. W., and Ungerleider, C. *The effects of role exchange questioning on empathetic perceptiveness*. No. 1, March, 48-52.
- Kieren, D. K. *The child and educational research—For whose benefit?* (Guest editorial). No. 1, March, 1-3.
- Kieren, T. E., and Southwell, B. *The development in children and adolescents of the construct of rational numbers as operators*. No. 4, December, 234-247.
- Kirman, J. M., and Goldberg, J. *Student teacher telephone conferencing with satellite maps as a monitoring device*. No. 4, December, 275-283.
- Liedtke, W., and Stott, G. *Games and game settings for the preschool child*. No. 3, September, 182-191.
- MacIver, D. A. *Education in Atlantic Canada* (Essay Review). No. 4, December, 288-293.
- Mason, J. M., and Kendall, J. R. *Facilitating reading comprehension through text structure manipulation*. No. 2, June, 68-76.

- McArthur, J. *Teacher socialization: The first five years*. No. 4, December, 264-274.
- Nwankwo, J. I., and Ohikhena, T. O. *Perception of the administrative role of the principal as a factor in student conflict in Nigeria*. No. 3, September, 147-159.
- Panuto, B., and White, D. *Achievement and the prediction of achievement in English first and second language children*. No. 2, June, 61-67.
- Pereira-Mendoza, L. *Heuristic strategies utilized by high school students*. No. 4, December, 213-220.
- Schlosser, L., and Algozzine, B. *The disturbing child: He or she?* No. 1, March, 30-36.
- Schnell, R. L. *History of childhood as history of education: A review of approaches and sources (Perspective)*. No. 3, September, 192-203.
- Shaw, E. *The effects of attentional focus on graphic discrimination*. No. 2, June, 103-116.
- Travis, L. D. and White, W. B. *Experimental manipulation of the recall of narrative material by five-year-olds*. No. 3, September, 137-146.
- Walker, L. *Newfoundland dialect interference in fourth grade spelling*. No. 4, December, 221-233.
- Williams, D. M. *A study of moral education in Surrey, B.C. secondary schools*. No. 2, June, 89-102.
- Young, J. H. *The curriculum decision-making preferences of school personnel*. No. 1, March, 20-29.





# PUBLICATIONS

FACULTY OF EDUCATION,  
UNIVERSITY OF ALBERTA  
EDMONTON T6G 2G5

## **The Alberta Journal of Educational Research**

**AJER** is a quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education. Published in March, June, September and December; subscription is \$8 per year. Address communications to the Editor, *AJER*, 732 Education South.

## **Elements**

**Elements** endeavours to help educators translate theory and findings from current research into classroom practice. Each issue focuses on a specific subject area, topic or theme. Published monthly September through April; subscription is \$4 per year. Address communications to the Editor, *Elements*, Department of Elementary Education, 539 Education South.

## **Indian-Ed**

**Indian-Ed** publishes articles, presents abstracts of research and reports news items related to the education of native peoples. Published four times per year; subscription is \$3. Address communications to the Editor, *Indian-Ed*, Department of Educational Foundations, 5-109 Education North.

## **The Canadian Administrator**

The **CA** serves administrators at all levels of educational organizations. Articles selected for publication are intended to familiarize practitioners with the findings and implications of research, and with analyses of current policies and practices. Published eight times per year October through May; subscription is \$5 per year. Address communications to Editor, *CA*, Department of Educational Administration, 7-104 Education North.

# CONTEMPORARY EDUCATION

The NOW Journal  
of Teaching and Learning

CE: Publishes articles by  
teachers, professors and  
school administrators

CE: Reviews current books  
in the broad field of  
education

CE: Helps educators with their  
professional tasks

CE: Offers lively editorials

CE: Gives you regular columns  
on provocative topics,  
people

a. A Turning Point\*

b. Meandering—

M. Dale Baughman\*\*

\*Readers describe special moments in education.

\*\*Professor of Secondary Education, and Editor of CONTEMPORARY EDUCATION.

CONTEMPORARY EDUCATION, Statesman Towers  
Room 1005, Indiana State University  
Terre Haute, Indiana 47809

Please send the next 4 issues of CONTEMPORARY EDUCATION  
for \$7.00 (\$10.00 for foreign countries).

\_\_\_\_\_ Payment enclosed \_\_\_\_\_ Bill me

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_

State \_\_\_\_\_ Zip \_\_\_\_\_



## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.

